DERIVING AN INSULIN RESISTANCE SCORE IN YOUTH WITH TYPE 1 DIABETES MELLITUS BASED ON CLINICAL RISK FACTORS

by

Zhen Li

BA in Animal Science, South China Agricultural University, China, 2013

Submitted to the Graduate Faculty of

the Department of Biostatistics

Graduate School of Public Health in partial fulfillment

of the requirement for the degree of

Master of Science

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Zhen Li

It was defended on

April 20, 2015

and approved by

Ingrid M. Libman, MD, PhD, Assistant Professor, Pediatrics
Children's Hospital, University of Pittsburgh Medical Center

Ada O Youk, PhD, Associate Professor, Biostatistics
Graduate School of Public Health,
University of Pittsburgh

Thesis Advisor: Vincent C. Arena, PhD, Associate Professor, Biostatistics,
Graduate School of Public Health, University of Pittsburgh

Vincent C. Arena, PhD

DERIVING AN INSULIN RESISTANCE SCORE IN YOUTH WITH TYPE 1 DIABETES MELLITUS BASED ON CLINICAL RISK FACTORS

Zhen Li, M.S.

University of Pittsburgh, 2015

ABSTRACT

Diabetes is a complicated chronic disease, and it is categorized into type 1 diabetes (T1D) and type 2 diabetes (T2D). Insulin sensitivity (IS) is lower in adults and adolescents with T1D compared to normal people, and a lower IS in T1D has been showed to be associated with longer-term complications. T1D is prevalent in children. The aim of this project was derive an insulin sensitivity (IS) score in children with T1D using noninvasive clinical predictors.

From a sample of 60 children undergoing a euglycaemic-hyperinsulinaemic clamp study at Children hospital of Pittsburgh of UPMC, a linear regression model was derived using clinical and laboratory measurements to predict insulin sensitivity. Because of the limitations of the small dataset, overfitting was an issue. We used a machine learning technique called Cross-Validation to help select predictors and to assess the performance. Data management and analysis were done using SAS 9.4.

One set of models built with only clinical variables were called clinical models. The other models used both clinical and laboratory variables were called research models. Two different outcome variables measures IS, glucose disposal rate (GDR) and glucose disposal rate (GDR) divided by free insulin, were used. After selecting the best models and checking the assumptions, the best model to predict GDR contained diastolic blood pressure percentile, systolic blood pressure percentile, gender, waist circumference, and diabetes duration. When the dependent

variable was GDR divided by free insulin, predictors in the best model included DBP percentile, HbA1C at the study time, waist circumference, leptin, and adiponectin/ leptin. These models had much better performance for type 1 diabetes than these models from the literature.

Public Health Significance: Identifying an IS predictive model based on routinely gathered clinical measurements and laboratory value is a valuable alternative to the invasive euglycaemic-hyperinsulinaemic clamp study. The current gold standard of insulin sensitivity, euglycaemic-hyperinsulinaemic clamp, is an invasive intravenous study requiring fasting overnight hospital study. The model makes it practical to use in epidemiological and screening studies.

TABLE OF CONTENTS

PREFACE ..... **XII**

1.0 INTRODUCTION **1**

2.0 DATA MANAGEMENT **3**

3.0 PATIENT SELECTION **4**

4.0 VARIABLES **5**

4.1.1 Outcome variables **5**

4.1.1.1 Glucose disposal rate (GDR) **5**

4.1.1.2 Glucose disposal rate (GDR) divided by free insulin **5**

4.1.2 Predictor variables **6**

4.1.2.1 Adiposity **6**

4.1.2.2 BMI **6**

4.1.2.3 Waist/Height Ratio **7**

4.1.2.4 Central obesity **7**

4.1.2.5 Waist circumference **7**

4.1.2.6 Lipid **8**

4.1.2.7 Blood pressure **8**

4.1.2.8 HbA1C **9**

4.1.3 Research variable **9**

4.1.3.1 Adiponectin 9

4.1.3.2 Leptin 9

5.0 DESCRIPTIVE STATISTICS 11

5.1.1 Continuous variables 11

5.1.1.1 Method 11

5.1.1.2 Result for predictors 11

5.1.1.3 Result for GDR 13

5.1.2 Categorical variables 13

5.1.2.1 Method 13

5.1.2.2 Result 14

5.1.3 Age-dependent variables 15

5.1.3.1 Method 15

5.1.3.2 Result 16

6.0 CORRELATIONS 18

6.1.1 Predictors correlated with each other 18

6.1.1.1 Results 18

7.0 UNIVARIATE REGRESSIONS 21

7.1.1 Log (GDR) 21

7.1.2 Log (GDR divided by free insulin) 22

8.0 MULTIPLE REGRESSION MODEL 24

8.1.1 Possible methods 24

8.1.2 Cross-Validation 25

8.1.2.1 Introduction 26

8.1.2.2 Criteria 26

8.1.2.3 Example 27

9.0 RESULTS 29

9.1.1 Log (GDR) 30

9.1.1.1 Best model 30

9.1.1.2 Check assumptions 31

- **Existence & Independent 31**
- **Linearity 32**
- **Homoscedasticity 32**
- **Normality 33**
- **Collinearity 35**

9.1.1.3 Diagnostics 35

9.1.2 Log (GDR divided by free insulin) 40

9.1.2.1 Best model 40

9.1.2.2 Assumption 41

- **Linearity 41**
- **Homoscedasticity 42**
- **Collinearity 44**

9.1.2.3 Diagnostics 44

10.0 DISCUSSION AND CONCLUSION 48

APPENDIX: SAS PROGRAMMING 50

BIBLIOGRAPHY 105

LIST OF TABLES

Table 1. Descriptive statistics for continuous variables 12

Table 2. Descriptive statistics for dependent variables 13

Table 3 Descriptive statistics for categorical variables 14

Table 4 Testing difference in log (GDR) by age groups 16

Table 5 Testing difference for log (GDR divided by free insulin) by age groups 17

Table 6 Correlations among variables 18

Table 7 Correlations among different measurements of adiposity 19

Table 8 Univariate regression models for log (GDR) 21

Table 9 Univariate regression models for log (GDR divided by free insulin) 23

Table 10 Example 27

Table 11 The best practical model for log (GDR) 30

Table 12 The best research model for log (GDR divided by free insulin) 30

Table 13 VIF for practical model 35

Table 14 VIF for research model 35

Table 15 Criteria of diagnostics for practical model 36

Table 16 Criteria of diagnostics for research model 38

Table 17 The best practical model for log (GDR divided by free insulin) 40

Table 18 The best research model for log (GDR divided by free insulin) 40

Table 19 VIF for practical model..... 44

Table 20 VIF for research model..... 44

Table 21 Criteria of diagnostics for practical model 44

Table 22 Criteria of diagnostics for practical model 46

LIST OF FIGURES

Figure 1. Distributions of dependent variables 16

Figure 2 Selection criteria 28

Figure 3 Scatterplot matrix for predictors and log (GDR) 32

Figure 4 Residuals distribution for practical model 33

Figure 5 Residuals distribution for research model 33

Figure 6 QQ-Plot for practical model 34

Figure 7 QQ-Plot for research model 34

Figure 8 Graphs for diagnostics 37

Figure 9 Graphs of diagnostics for research model 39

Figure 10 Scatterplot matrix for predictors and log (GDR divided by free insulin), ... 41

Figure 11 Residuals distribution for practical model 42

Figure 12 Residuals distribution for research model 42

Figure 13 QQ-Plot for practical model 43

Figure 14 QQ-Plot for research model 43

Figure 15 Graphs of diagnostics for practical model 45

Figure 16 Graphs of diagnostics for research model 47

PREFACE

Thank you very much to my thesis advisor, Dr. Vincent Arena, for providing me the opportunity to participate in this project. All analysis and programming were finished by myself under your guidance.

Clinical interpretation was done under instruction of Dr. Ingrid M. Libman, and thanks to her efforts, which gave more clinical meaning to the result. Also, Thank you for sharing your dataset with me.

Thanks to Dr. Ada Youk. Dr. Youk' s excellent regression course helped me a lot when performing analysis.

We acknowledge partial support for this project by National Institutes of Health (NIH) grants, NHLBI K23HL085287 Award, principal investigator Dr. Ingrid Libman.

1.0 INTRODUCTION

Diabetes is a complicated disease with two different types. Type 2 diabetes (T2D) is prevalent in adults, whereas type 1 diabetes (T1D) is more prominent in juveniles [1] T1D is becoming more and more prevalent, and its incidence rate increases about 3% each year. T1D is present when a child's pancreas cannot produce a sufficient amount of insulin [2]. A child needs to replace the missing insulin [3]. Regular blood sugar monitoring and insulin delivery can improve the management of T1D in children [3]. Because of the above reasons, T1D in children is known as juvenile diabetes or insulin-dependent diabetes [3].

Insulin is a hormone made by the pancreas. Insulin is a hormone that makes cells absorb blood sugar for energy or storage. [4]. For example, when you eat a meal, the blood glucose increases after digesting food, when insulin takes action in order to store the glucose in muscle as glycogen or as fat [5]. For example, lower IS in T1D causes a more atherogenic lipid profile, increased carotid intima-media thickness (IMT) [6], and increased risk of complications (coronary artery calcifications and albuminuria) [7].

Methods to measure or surrogate the insulin sensitivity are HOMA-insulin resistant (HOM-IR), the quantitative insulin sensitivity check index (QUICKI) [8], the SEARCH IS score [9], and euglycaemic-hyperinsulinaemic clamp [10]. The first two methods, HOM-IR and QUICKI, cannot be applied on insulin-treated patients and may be inaccurate in youths with type I diabetes, because the assumptions of these methods are normal glucose levels and preserved

insulin and C-peptide secretion [10]. SEARCH [10] IS was derived from individuals with type 1 and type 2 diabetes. The euglycaemic-hyperinsulinaemic clamp, which is the gold standard, is an invasive intravenous procedure requiring a fasting overnight hospital stay. This is not applicable for larger epidemiological studies.

The aim of this study is to investigate in youth with T1D if clinical variables can estimate IS as determined by the gold standard hyperinsulinemic-euglycemic clamp. This thesis, data management and analysis will be performed.

This study was conducted at Children Hospital of Pittsburgh, and children from 12 years to 18 years of age were recruited.

The study was approved by the University of Pittsburgh Institutional Review Board and all subjects were consented priority for patients.

2.0 DATA MANAGEMENT

The original data were collected three years ago, so updating the data and incorporating new information was necessary. Because some variables were calculated by hand, these variables were recalculated, such as age by years and diabetes duration by years. Some new variables were created based on predictors, such as waist to height ratio, central obesity, weight to height ratio, adiponetin to leptin ratio, waist percentage, waist category, and exogenous-glucose / free insulin. If the waist-height ratio was greater than 0.5, then central obesity was coded as 1. Otherwise, it was coded as 0. If the waist percentage was smaller than 85%, then waist category was coded as 1. Otherwise, it was given the value of 2. For the multiple-level categorical variable waist percentage continuous values were created. Waist percentage was ordered by codes 0, 1, 2, 3, 4, 5 and 6, and its correspond continuous values that centered around were 5, 17.5, 37.5, 62.5, 80, 87.5 and 95. All the calculations were done in SAS 9.4.

A review of all data was performed and all missing data was brought to the attention. Confirmation of missing data was also made, and no new data was added. All of the programming was completed in SAS 9.4.

3.0 PATIENT SELECTION

Subjects were eligible based on the following criteria: diagnosed with autoimmune T1DM with presence of at least one β-cell antibody at diagnosis, ages between 12 and 18 years, pubertal stage Tanners at II to V, duration of diabetes greater than 2 years, and HbA1c lower than 10%. Participants were admitted to the Pediatric Clinical and Translational Research Center (PCTRC) for an overnight stay.

A 3-hour hyperinsulinemic-euglycemic clamp was performed to measure in vivo insulin sensitivity. An intravenous catheter was placed in an antecubital vein that continuously infused insulin and glucose in order to achieve endogenous stable insulin concentrations. At the same time, another catheter was placed retrograde in a dorsal vein of the contralateral hand to draw steady-state blood samples.

All subjects were studied following a 10-hour overnight fasting period. All of these visits were completed within 24 hours, and the subjects were discharged after lunch. Among all of the eligible subjects, 60 had complete data, so their information was used to perform this analysis.

4.0 VARIABLES

Originally, there were 108 variables and two outcome variables in the dataset. After selecting variables using clinical knowledge, 18 variables were considered as potential predictors to build the models.

4.1.1 Outcome variables

4.1.1.1 Glucose disposal rate (GDR)

This is a clamp-derived measure of insulin sensitivity.

4.1.1.2 Glucose disposal rate (GDR) divided by free insulin

GDR divided by free insulin is an adjusted variable that takes into account the insulin levels during the clamp.

4.1.2 Predictor variables

4.1.2.1 Adiposity

Adiposity, which is the state of obese, is an important indicator for cardio-metabolic disorders. In addition, it can help people predict their health conditions. There are many different measurements, but researchers have no agreement on the best method to measure adiposity.

4.1.2.2 BMI

BMI (Body Mass Index) measures body fat based on height and weight (BMI=weight/height^2), and can be used as a surrogate indicator of cardiovascular risk [11]. BMI indicates whether or not you are thin or fat. For example, the higher your BMI, the more fat you have compared to your height. Absolute value of BMI is not suitable for children, because the body frame changes over time, and children of different ages and sex have different standards based on BMI.

Two alternatives are BMI Z Score and BMI percentile. BMI z score, a standardized value, is more meaningful for a statistician, but it is sometimes harder to understand for physicians. The z score is the stander value from a N (0,1) distribution for children of a specific age, sex and race group. BMI percentile is easier for a non-statistician, such as patients and physicians, to understand. Because the purpose of this project was to come up with an interpretable model, we used the BMI percentile.

4.1.2.3 Waist/Height Ratio

Some researchers believe that waist to height ratio is a more appropriate way than BMI to measure whether or not you are overweight. This measurement is valid for short and tall people [12], and it is valid for both children and adults.

4.1.2.4 Central obesity

Using BMI to define obesity does not consider different body frames, and it does not differentiate between muscles and fat. Different body frames can predict human health. For example, the accumulation of belly fat is associated with several diseases, such as a high diabetes risk, high blood pressure and cardiovascular disease, and central obesity (CO). CO is defined as waist to height ratio which is greater than 0.5.

4.1.2.5 Waist circumference

In the past, investigators considered using waist circumference to measure people's health conditions. Waist circumference (WC) is an important clinical measurement. For example, it can significantly predict cardiovascular disease, and it can explain obesity-related health risks as well. Previous reports demonstrated that WC is a crucial predictor of abdominal fat and metabolic dysfunction. Researchers reported that WC itself could explain the 59% variance of Glucose disposal rate (GDR) [10]. Although WC is crucial in diabetes projects, standards for waist circumference change for children over age. In other words, there is no consistent standard of WC for children to indicate their health conditions. Because of these reasons, we consider using waist percentile instead. There are potential choices include how to treat WC as a continuous or categorical variable. When building models, we tried different forms of WC, including a continuous WC variable using percentiles, an ordered categorical WC variable using

7 levels, and a categorical WC variable using 2 levels. One of the goals of this project was to find out a WC form that might improve the performance of the model.

4.1.2.6 Lipid

Numerous studies have shown that obese children have a more adverse lipid profile than non-obese children [13]. T1D patients always have lipid disorders [14]. For T1D patients with optimal glycaemic control, low-density lipoprotein (LDL) cholesterol and plasma triglycerides are slightly decreased, whereas high-density lipoprotein (HDL) cholesterol is slightly increased. Insulin plays a crucial role in regulating lipid metabolism [14].

4.1.2.7 Blood pressure

In the general population, high blood pressure is more common in people with diabetes [15]. According to a UK report, around 3 in 10 people with T1D and 8 in 10 people with T2D develop high blood pressure at some stage [15]. For diabetes patients, controlling their blood pressure is very important, because high blood pressure is a significant risk factor that can increase disease risks, such as heart disease and cardiovascular disease. For patients, lowering their blood pressure by taking medication is crucial. Systolic blood pressure (SBP) is the blood pressure in the arteries when the heart contracts. Diastolic blood pressure (DBP) is the blood pressure in the arteries between each heartbeat. A normal SBP is lower or equal to 120, and if above this level, the risk of developing heart disease will increase. A normal SBP is lower or equal to 80, and if above this level, the patients will be considered to have hypertensive or have high blood pressure [16]. However, both of these standards are for adults, not for children. Because our project was for diabetic children, SBP z score and DBP z score, or SBP percentile and DBP percentile were considered.

4.1.2.8 HbA1C

HbA1C can be referred to as hemoglobin A1c or A1c. Hemoglobin is a protein within red blood cells that carries oxygen throughout a person's body. HbA1c develops when hemoglobin joins with glucose in the blood. HbA1C can provide information about a person's average levels of blood glucose, which is called blood sugar [17]. Diabetes patients have A1C tested at least twice a year, and this test can show if their diabetes plans are working. For people with diabetes, the higher the HbA1c, the greater the risk of developing diabetes-related complications [18].

4.1.3 Research variable

4.1.3.1 Adiponectin

Adiponectin is a novel adipocyte-specific protein. According to reports, it plays a role in developing insulin resistance, and it can affect the body's response to insulin. Adiponectin is lower in obese people than in lean people [18]. On the other hand, measuring adiponectin is not usually done, and in order to make our model more generalized, we considered using adiponectin later.

4.1.3.2 Leptin

Leptin helps regulate body fat. Obese people tend to have higher leptin levels than normal weight individuals. According to publications, treatment with leptin reverses hyperglycemia in animal models with poorly controlled T1D and T2D [19]. The leptin hormone regulates appetite, metabolism, and body weight. Some researchers believe leptin can help to discover new therapies to control hyperglycemia in patients with T1D and T2D [19]. Like

adipnectin, it is unusual to measure leptin in the standard when diagnosing diabetic patients, and it is typically a research measure.

5.0 DESCRIPTIVE STATISTICS

5.1.1 Continuous variables

5.1.1.1 Method

The analysis was based on 60 eligible subjects and 20 potential continuous predictors. Descriptive statistics include number of available value, mean, standard error, median, 5%, 25%, 75%, 95%, minimum and maximum. This step provided a big picture about the dataset for researchers, and it was also a good way to check the quality of the data.

5.1.1.2 Result for predictors

Table 1. Descriptive statistics for continuous variables

Predictor	N	Mean	Standard Error	Median	5%	25%	75%	95%	Min	Max
Age (years)	60	15.042	0.232	14.833	12.417	13.492	16.625	17.792	12.250	17.917
Diabetes Duration (years)	60	6.928	0.415	6.917	2.500	4.250	8.792	12.917	2.417	14.750
Diastolic Blood Pressure	60	66.559	0.815	66.500	55.500	62.330	70.500	77.500	52.000	79.000
Diastolic Blood Pressure (%)	60	53.663	2.570	53.469	16.720	43.040	70.074	84.607	7.779	92.268
Systolic Blood Pressure	60	116.727	1.146	116.500	103.500	111.800	121.000	133.500	97.000	143.000
Systolic Blood Pressure (%)	60	61.605	3.099	63.694	19.763	48.745	79.550	96.981	4.338	99.439
Cholesterol	60	159.750	3.715	161.000	106.000	141.000	179.500	207.000	92.000	227.000
High-density lipoprotein	60	53.830	1.473	53.800	38.200	44.600	62.350	72.950	34.000	73.900
Low-density lipoprotein	60	90.248	3.006	89.900	50.950	72.900	106.500	132.400	37.600	140.000
Triglyceride	60	78.450	3.914	75.500	45.500	58.000	89.500	131.000	36.000	203.000
BMI (%)	60	77.899	2.878	86.718	23.496	74.028	93.334	97.500	12.244	99.165
Waist	59	77.695	1.239	76.000	67.000	70.000	84.000	99.000	58.000	100.000
Waist/Height	59	0.471	0.008	0.452	0.370	0.433	0.512	0.600	0.368	0.618
Waist (%, midpoint)	59	55.212	3.721	62.500	5.000	37.500	80.000	95.000	5.000	95.000
Triglyceride/HDL	60	1.552	0.108	1.336	0.763	1.097	1.729	3.263	0.523	5.273
Adiponectin/leptin	60	1.595	0.254	0.764	0.144	0.361	2.244	6.967	0.088	8.889
Adiponectin	60	9.660	0.453	9.700	4.600	7.250	11.600	15.100	3.100	20.300
Leptin	60	16.178	1.810	12.250	2.100	4.050	26.100	43.450	1.500	56.800
HbA1C (at the time of study)	60	8.053	0.113	7.950	6.700	7.400	8.600	9.700	6.300	9.900
HbA1C (Last 5 clinic visits)	60	8.057	0.122	7.960	6.630	7.420	8.460	9.790	6.480	10.500

The age of the patients was 15.0 ± 0.2 years, and ranged between 12.3 and 17.9. The diabetes duration was 6.9 ± 0.4 years, and ranged between 2.4 and 14.8. These results were consistent with our project design. The median BMI percentile was 86.718%, and this indicated that over 50% of the children in our project were over-weight. The DBP was 6.9 % ± 0.4 %. The SBP was 116.7 % ± 1.1 %. The HbA1c at the time of study did not change much from the HbA1c during the last 5 clinical visits.

5.1.1.3 Result for GDR

Table 2. Descriptive statistics for dependent variables

Predictor	N	Mean	Standard Error	Median	5%	25%	75%	95%	Min	Max
GDR	60	8.149	0.376	7.495	4.195	6.235	10.105	14.155	3.677	15.450
GDR divided by free insulin	59	8.760	0.464	8.059	3.605	6.175	10.454	15.731	2.852	18.417

The value of GDR was 8.1 ± 0.4 mg/kg/min, and ranged from 3.7 to 15.5 mg/kg/min. The value of GDR divided by free insulin based on 59 subjects did not differ appreciable from GDR. The value of GDR divided by free insulin was 8.8 ± 0.5 mg/kg/min/pmol/l, and ranged from 2.9 to 18.4 mg/kg/min/pmol/l.

5.1.2 Categorical variables

5.1.2.1 Method

Frequency analysis was performed on the 60 eligible subjects and included 5 categorical variables. Waist circumference defined as a categorical variable with values 0, 1, 2, 3, 4, 5 and 6, represented percentages 0%~10%, 10%~25%, 25%~50%, 50%~75%, 75%~85%, 85%~90% and

90%~100% for each category. We also coded WC as a dichotomous variable using 85%. Tanner stage describes puberty stage. For the tanner stage variable, if the subject was a girl, then the information of tanner_stage_breast was the value of the tanner stage variable. Otherwise, the variable tanner_stage_pubic_hair was used. The value of the tanner stage was coded as 1 if its original value was less than 5. Otherwise, the tanner stage was coded as 2 if its original value was 5. Central obesity is based on the value of waist to height ratio, and the cut-point is 0.5.

5.1.2.2 Result

Table 3 Descriptive statistics for categorical variables

Gender	Frequency	Percent
Male	34	56.67
Female	26	43.33

Central obesity	Frequency	Percent
Yes	18	30.51
No	41	69.49
Frequency Missing = 1		

Waist	Frequency	Percent
<85%	48	81.36
>=85%	11	18.64
Frequency Missing = 1		

Tanner stage	Frequency	Percent
II, III, IV	22	36.67
V	38	63.33

Waist percentage	Frequency	Percent
0%	4	6.78
10%	10	16.95
25%	8	13.56
50%	19	32.20
75%	7	11.86
85%	4	6.78
90%	7	11.86
Frequency Missing = 1		

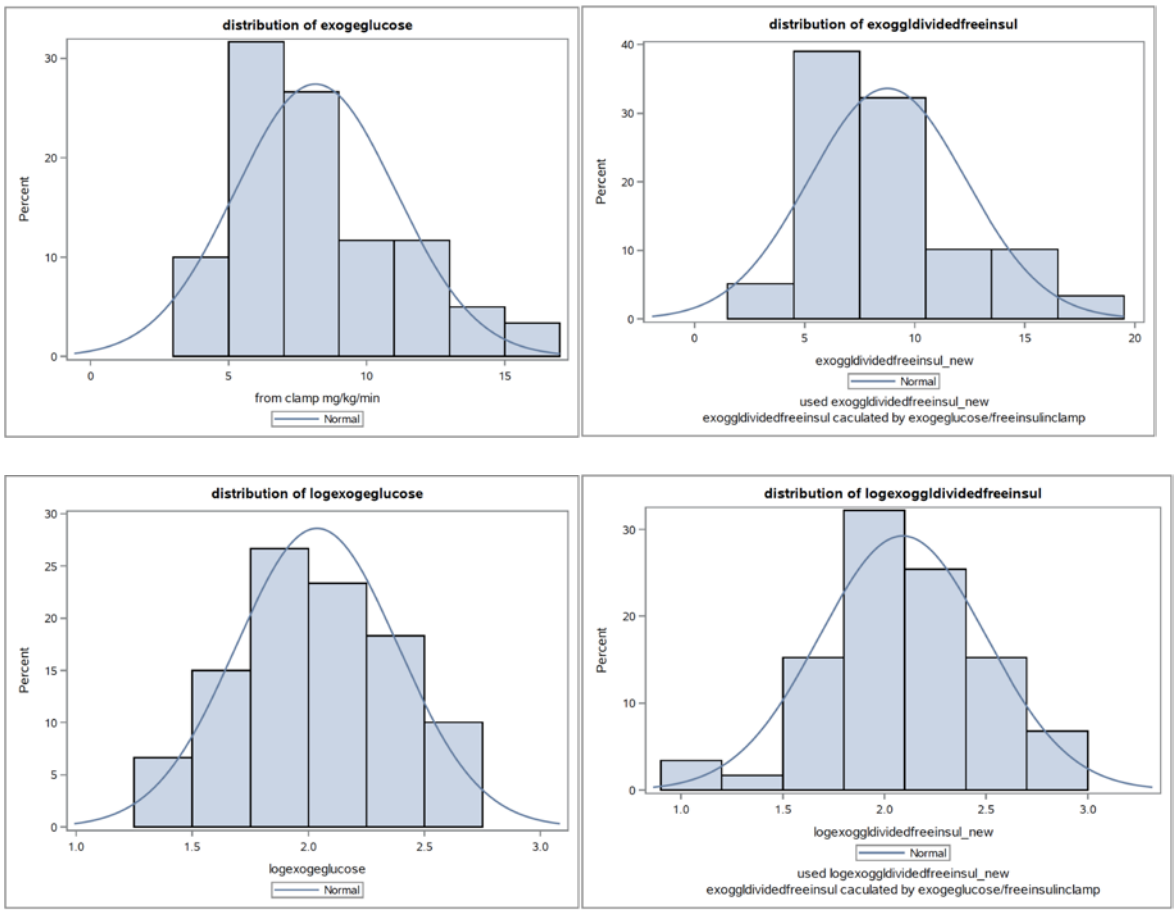
There were 34 boys and 26 girls in our analysis. Eighteen were central obesity problem, and 41 did not. Eleven of the 60 patients had a waist percentile greater than or equal to 85%. Thirty-eight children were post puberty.

5.1.3 Age-dependent variables

GDR and GDR divided by free insulin are not constant over time, and may be age related. For boys, GDR is believed to decrease from 12 to 16 years old, and after 16 years old, GDR increases. On the other hand, for girls, GDR decreases from 11 to 15 years old, and after 15 years old, GDR increases.

5.1.3.1 Method

Analyses were separately conducted creating descriptive statistics stratified by age and genders. T-tests were used to formally test the difference between different ages. The distribution of GDR and GDR divided by free insulin, which indicates that they were not normally distributed as seen in Figure 1. According to the theory of the linear model, normally distributed dependent variables usually have residual that are normally distributed. The normal distribution of residuals is one of the assumptions of linear regression analysis. Therefore, a normally distributed dependent variable can improve the performance of the regression model. For this reason, log transformation was used. In addition, the assumption of the t-test is normality, so log transformation was necessary when conducting t-tests. The distribution graphs of log (GDR) and log (GDR divided by free insulin) demonstrated that they were normally distributed after transforming. Thus t-tests were performed using log (GDR) and log (GDR divided by free insulin) separately.



Note: GDR=exogenous-glucose

Figure 1. Distributions of dependent variables

5.1.3.2 Result

Table 4 Testing difference in log (GDR) by age groups

Age category	N	Mean	Standard Error	Median	5%	25%	75%	95%	Min	Max
Girl: 11-15 or Boy: 12-16	38	2.097	0.050	2.044	1.605	1.878	2.399	2.660	1.472	2.725
Girl: >=15 or Boy: >=16	22	1.936	0.084	1.941	1.348	1.681	2.235	2.631	1.302	2.738
F-test (Equal of Variance): P-value= 0.1933; Equal variance t-test: P-value = 0.0836;										

Table 5 Testing difference for log (GDR divided by free insulin) by age groups

Age category	N	Mean	Standard Error	Median	5%	25%	75%	95%	Min	Max
Girl: 11-15 or Boy: 12-16	37	2.153	0.063	2.133	1.542	1.893	2.376	2.756	1.125	2.913
Girl: >=15 or Boy: >=16	22	1.985	0.093	1.952	1.282	1.777	2.254	2.742	1.048	2.838
F-test (Equal of Variance): P-value= 0.5209; Equal variance t-test: P-value = 0.1280;										

The F-test for equality of variance was not statistically significant (P-value = 0.193) indicating that there was equal variance for log (GDR), so equal variance t-test was most appropriate. Because the P-value of the t-test was 0.0836, there was no statistically significant difference in log (GDR) between the two different age groups. Similarly, the F-test (P-value = 0.521) showed equal variance for log (GDR divided by free insulin), so the equal variance t-test was used. Because the P-value of the t-test was 0.1280, there was no statistically significant difference of log (GDR divided by free insulin) between the two different age groups.

6.0 CORRELATIONS

6.1.1 Predictors correlated with each other

Linear correlations between all the predictors were performed. The goal of this step was to preliminarily investigate any collinearity problem in the dataset, and the results can provide insight to the investigator about the relationship between predictors.

6.1.1.1 Results

Table 6 Correlations among variables

	HbA1C (at the time of study)	HbA1C (at last 5 clinic visits)	Cholesterol	Triglyceride	HDL	LDL	Age (year)	Waist circumference	BMI (%)
HbA1C (at the time of study)	1.000	0.777 <.0001	0.022 0.867	-0.115 0.375	-0.054 0.679	0.081 0.530	-0.130 0.315	0.027 0.841	-0.165 0.197
HbA1C (at last 5 clinic visits)	0.777 <.0001	1.000	0.070 0.590	-0.055 0.674	-0.012 0.929	0.104 0.421	-0.092 0.476	-0.037 0.778	-0.071 0.582
Cholesterol	0.022 0.867	0.070 0.590	1.000	0.459 0.0002	0.426 0.0006	0.916 <.0001	-0.077 0.551	0.138 0.294	0.212 0.099
Triglyceride	-0.115 0.375	-0.055 0.674	0.459 0.0002	1.000	-0.162 0.208	0.390 0.002	0.144 0.266	0.377 0.003	0.260 0.041
HDL	-0.054 0.679	-0.012 0.929	0.426 0.0006	-0.162 0.208	1.000	0.079 0.540	-0.122 0.344	-0.460 0.0002	-0.333 0.008
LDL	0.081 0.530	0.104 0.421	0.916 <.0001	0.390 0.002	0.079 0.540	1.000	-0.075 0.565	0.298 0.021	0.361 0.004
Age (year)	-0.130 0.315	-0.092 0.476	-0.077 0.551	0.144 0.266	-0.122 0.344	-0.075 0.565	1.000	0.175 0.180	0.130 0.313
Waist circumference	0.027 0.841	-0.037 0.778	0.138 0.294	0.377 0.003	-0.460 0.0002	0.298 0.021	0.175 0.180	1.000	0.657 <.0001
BMI (%)	-0.165 0.197	-0.071 0.582	0.212 0.099	0.260 0.041	-0.333 0.008	0.361 0.004	0.130 0.313	0.657 <.0001	1.000
Systolic BP (%)	0.023 0.860	0.057 0.660	0.230 0.073	0.069 0.594	0.175 0.173	0.183 0.155	-0.249 0.051	-0.117 0.372	0.029 0.821
Diastolic BP (%)	0.150 0.241	0.120 0.349	0.269 0.035	0.020 0.879	0.213 0.096	0.224 0.080	-0.129 0.318	0.073 0.581	-0.051 0.694
Leptin	0.033 0.801	-0.005 0.972	0.448 0.0003	0.097 0.455	0.219 0.087	0.423 0.001	0.012 0.928	0.220 0.092	0.281 0.027
Adiponectin	0.002 0.988	0.026 0.838	-0.011 0.933	-0.091 0.482	0.303 0.017	-0.137 0.287	0.005 0.970	-0.311 0.016	-0.289 0.022
Waist/height	0.073 0.581	0.006 0.963	0.187 0.154	0.363 0.004	-0.389 0.002	0.327 0.011	-0.069 0.603	0.897 <.0001	0.559 <.0001
Triglyceride/HDL	-0.101 0.435	-0.066 0.612	0.194 0.131	0.904 <.0001	-0.535 <.0001	0.269 0.034	0.144 0.264	0.522 <.0001	0.339 0.007
Waist (%)	0.054 0.680	0.001 0.993	0.117 0.375	0.225 0.084	-0.397 0.002	0.281 0.030	-0.265 0.041	0.826 <.0001	0.629 <.0001
Adiponectin/leptin	-0.193 0.134	-0.124 0.336	-0.305 0.016	-0.214 0.095	0.206 0.109	-0.424 0.001	0.035 0.784	-0.381 0.003	-0.526 <.0001
Duration (year)	0.123 0.337	0.206 0.105	-0.050 0.702	0.062 0.635	-0.106 0.413	-0.024 0.85	0.257 0.044	0.171 0.191	0.105 0.414

Table 6 Continued

	Systolic BP (%)	Diastolic BP (%)	Leptin	Adiponectin	Waist/ height	Triglyceride /HDL	Waist (%)	Adiponectin /leptin	Duration (year)
HbA1C (at the time of study)	0.023 0.860	0.150 0.241	0.033 0.801	0.002 0.988	0.073 0.581	-0.101 0.435	0.054 0.680	-0.193 0.134	0.123 0.337
HbA1C (at last 5 clinic visits)	0.057 0.660	0.120 0.349	-0.005 0.972	0.026 0.838	0.006 0.963	-0.066 0.612	0.001 0.993	-0.124 0.336	0.206 0.105
Cholesterol	0.230 0.073	0.269 0.035	0.448 0.0003	-0.011 0.933	0.187 0.154	0.194 0.131	0.117 0.375	-0.305 0.016	-0.050 0.702
Triglyceride	0.069 0.594	0.020 0.879	0.097 0.455	-0.091 0.482	0.363 0.004	0.904 <.0001	0.225 0.084	-0.214 0.095	0.062 0.635
HDL	0.175 0.173	0.213 0.096	0.219 0.087	0.303 0.017	-0.389 0.002	-0.535 <.0001	-0.397 0.002	0.206 0.109	-0.106 0.413
LDL	0.183 0.155	0.224 0.080	0.423 0.001	-0.137 0.287	0.327 0.011	0.269 0.034	0.281 0.030	-0.424 0.001	-0.024 0.85
Age (year)	-0.249 0.051	-0.129 0.318	0.012 0.928	0.005 0.970	-0.069 0.603	0.144 0.264	-0.265 0.041	0.035 0.784	0.257 0.044
Waist circumference	-0.117 0.372	0.073 0.581	0.220 0.092	-0.311 0.016	0.897 <.0001	0.522 <.0001	0.826 <.0001	-0.381 0.003	0.171 0.191
BMI (%)	0.029 0.821	-0.051 0.694	0.281 0.027	-0.289 0.022	0.559 <.0001	0.339 0.007	0.629 <.0001	-0.526 <.0001	0.105 0.414
Systolic BP (%)	1.000	0.515 <.0001	0.077 0.553	-0.009 0.947	0.0002 0.999	0.009 0.946	-0.031 0.813	-0.100 0.439	0.079 0.540
Diastolic BP (%)	0.515 <.0001	1.000	0.353 0.005	-0.052 0.687	0.150 0.254	-0.056 0.665	0.103 0.435	-0.200 0.119	-0.093 0.471
Leptin	0.077 0.553	0.353 0.005	1.000	-0.080 0.539	0.309 0.016	-0.053 0.683	0.266 0.040	-0.596 <.0001	0.093 0.473
Adiponectin	-0.009 0.947	-0.052 0.687	-0.080 0.539	1.000	-0.338 0.008	-0.202 0.116	-0.281 0.030	0.464 0.0001	-0.105 0.413
Waist/height	0.0002 0.999	0.150 0.254	0.309 0.016	-0.338 0.008	1.000	0.479 0.0001	0.858 <.0001	-0.481 0.0001	0.100 0.447
Triglyceride/HDL	0.009 0.946	-0.056 0.665	-0.053 0.683	-0.202 0.116	0.479 0.0001	1.000	0.360 0.005	-0.205 0.109	0.084 0.514
Waist (%)	-0.031 0.813	0.103 0.435	0.266 0.040	-0.281 0.030	0.858 <.0001	0.360 0.005	1.000	-0.466 0.0002	0.063 0.635
Adiponectin/leptin	-0.100 0.439	-0.200 0.119	-0.596 <.0001	0.464 0.0001	-0.481 0.0001	-0.205 0.109	-0.466 0.0002	1.000	-0.173 0.179
Duration (year)	0.079 0.540	-0.093 0.471	0.093 0.473	-0.105 0.413	0.100 0.447	0.084 0.514	0.063 0.635	-0.173 0.179	1.000

Table 7 Correlations among different measurements of adiposity

	Waist circumference	BMI (%)	Waist/height	Waist (%)
Waist circumference	1.000	0.657 <.0001	0.897 <.0001	0.826 <.0001
BMI (%)	0.657 <.0001	1.000	0.559 <.0001	0.629 <.0001
Waist/height	0.897 <.0001	0.559 <.0001	1.000	0.858 <.0001
Waist (%)	0.826 <.0001	0.629 <.0001	0.858 <.0001	1.000

The correlation coefficient matrix indicated that average HbA1c at the last 5 clinic visits highly correlated with HbA1c at the time of study ($r = 0.777$, $p < .0001$), which is somewhat intuitive. Because HbA1c at the time of study was more meaningful for clinics, we preferred to use this measure.

LDL and cholesterol had high correlation ($r = 0.916$, $p < .0001$). The correlation between LDL ($r=0.143$, $p < 0.003$) and log (GDR) was higher than the correlation between cholesterol

(r=0.084, p = 0.025) and log (GDR). When the dependent variable was log (GDR divided by free insulin), the correlation between cholesterol (p = 0.011) and log (GDR) was not significant, but the correlation between LDL (p = 0.076) and log (GDR) was significant.

Triglyceride was highly correlated with triglyceride hdl ratio (r= 0.904, p < .0001). We preferred to use triglyceride hdl ratio rather than triglyceride because log (GDR) had higher correlation with triglyceride hdl ratio (r=0.094, p = 0.017) than triglyceride (r=0.085, p = 0.024). Another reason why we used triglyceride hdl ratio was it contained more information than triglyceride. Neither triglyceride hdl ratio nor triglyceride had significantly statistical relationship with log (GDR divided by free insulin).

BMI percentile, waist circumference, waist to height ratio, and waist percentile had high correlations with each other. All four of these predictors were important to describe adiposity. In order to find the predictor, which can contribute most to the models comparing with other adiposity measurements, we separately considered these four variables when building models.

7.0 UNIVARIATE REGRESSIONS

7.1.1 Log (GDR)

Regression models were built using each potential predictor and log (GDR) to investigate the relationships between each predictor and the dependent variable.

Table 8 Univariate regression models for log (GDR)

	Predictor	N	β	Standard Error	Standard -ized β	P-Value	R-Square	Adjusted R-Square
Log (GDR)	Age (year)	58	-0.02458	0.02537	-0.12620	0.3366	0.0159	-0.0010
	Diastolic Blood Pressure (%)	60	-0.00617	0.00215	-0.35247	0.0057	0.1242	0.1091
	Systolic Blood Pressure (%)	60	-0.00240	0.00188	-0.16502	0.2077	0.0272	0.0105
	Cholesterol	60	-0.00351	0.00152	-0.28969	0.0248	0.0839	0.0681
	High-density lipoprotein	60	0.00596	0.00393	0.19512	0.1352	0.0381	0.0215
	Low-density lipoprotein	60	-0.00565	0.00182	-3.7756	0.0029	0.1426	0.1278
	Triglyceride	60	-0.00335	0.00144	-0.29093	0.0241	0.0846	0.0689
	BMI (%)	60	-0.00701	0.00184	-0.44824	0.0003	0.2009	0.1871
	Waist	59	-0.02040	0.00408	-0.55207	≤.0001	0.3048	0.2926
	Waist/Height	59	-3.12238	0.59250	-0.57236	≤.0001	0.3276	0.3158
	Waist (%)	59	-0.00628	0.00134	-0.52576	≤.0001	0.2764	0.2637
	Triglyceride /HDL	60	-0.12789	0.05213	-0.30663	0.0172	0.0940	0.0784
	Adiponectin	60	0.03161	0.01237	0.31809	0.0133	0.1012	0.0857
	Leptin	60	-0.01231	0.00284	-0.49499	<0.001	0.2450	0.2320
	Adiponectin/leptin	60	0.10267	0.01895	0.57978	<0.0001	0.3361	0.3247
	Last_HbA1C	60	-0.05335	0.05164	-0.13442	0.3059	0.0181	0.0011
	Avrg_HbA1C	60	-0.03234	0.04835	-0.08750	0.5062	0.0077	-0.0095
	Gender	60	-0.27190	0.08437	-0.38972	0.0021	0.1519	0.1373
	Waist (category: 7 levels)	59	-0.11087	0.02293	-0.53940	<0.0001	0.2909	0.2785
	Central obesity	59	-0.36097	0.08816	-0.47674	0.0001	0.2273	0.2137
	Waist (category: 2 levels)	59	-0.38184	0.10724	-0.42656	0.0008	0.1820	0.1676
	Age (category: 2 levels)	60	-0.16157	0.09178	-0.22520	0.0836	0.0507	0.0343
	Diabetes duration (year)	60	-0.02618	0.01382	-0.24145	0.0631	0.0583	0.0421

From the results of the univariate regression models, we found that log (GDR) significantly related with DBP, cholesterol, LDL, triglyceride, BMI, waist circumference, waist/height ratio, waist percentile, triglyceride/HDL ratio, adiponectin/leptin ratio, gender, waist circumference (7 categories), waist circumference (2 categories), and central obesity. Among these predictors, waist/height ratio had the highest relationship with log (GDR) ($r= 0.328$, $p = < 0.0001$). In addition, waist circumstance (2 categories) ($r= 0.182$, $p = 0.0008$), central obesity ($r= 0.223$, $p = 0.0001$) and female ($r= 0.152$, $p = 0.0021$) also had high correlation with log (GDR).

7.1.2 Log (GDR divided by free insulin)

Another dependent variable that we were interested in was log (GDR by free insulin). In order to know the relationships between each predictor and the dependent variable, univariate regression models were built.

Table 9 Univariate regression models for log (GDR divided by free insulin)

	Predictor	N	β	Standard Error	Standard -ized β	P-Value	R-Square	Adjusted R-Square
Log (GDR divided free insulin)	Age (year)	58	-0.02241	0.03032	-0.09741	0.4630	0.0095	-0.0079
	Diastolic Blood Pressure (%)	59	-0.00643	0.00261	-0.30983	0.0169	0.0960	0.0801
	Systolic Blood Pressure (%)	59	-0.00038	0.00230	-0.02175	0.8701	0.0005	-0.0171
	Cholesterol	59	-0.00333	0.00184	-0.23317	0.0755	0.0544	0.0378
	High-density lipoprotein	59	0.00629	0.00466	0.17622	0.1818	0.0311	0.0141
	Low-density lipoprotein	59	-0.00580	0.00220	-0.33020	0.0106	0.1090	0.0934
	Triglyceride	59	-0.00221	0.00175	-0.16478	0.2123	0.0272	0.0101
	BMI (%)	59	-0.00881	0.00212	-0.48156	0.0001	0.2319	0.2184
	Waist	58	-0.02295	0.00486	-0.53356	<.0001	0.2847	0.2719
	Waist/Height	58	-3.39827	0.71811	-0.53447	<.0001	0.2857	0.2729
	Waist (%)	58	-0.00702	0.00161	-0.50310	<.0001	0.2531	0.2398
	Triglyceride /HDL	59	-0.09150	0.06317	-0.18843	0.1529	0.0355	0.0186
	Adiponectin	59	0.04012	0.01463	0.34144	0.0081	0.1166	0.1011
	Leptin	59	-0.01503	0.00328	-0.51895	<.0001	0.2693	0.2565
	Adiponectin/leptin	59	0.12234	0.02207	0.59190	<.0001	0.3504	0.3390
	HbA1C (Last 5 clinic visits)	59	-0.09032	0.06013	-0.19512	0.1386	0.0381	0.0212
	HbA1C (at the time of study)	59	-0.08813	0.05698	-0.20071	0.1274	0.0403	0.0234
	Gender	59	-0.29020	0.10119	-0.35511	0.0058	0.1261	0.1108
	Waist (category: 7 levels)	58	-0.12118	0.02765	-0.50531	<.0001	0.2553	0.2420
	Central obesity	58	-0.46451	0.10273	-0.51716	<.0001	0.2675	0.2544
	Waist (category: 2 levels)	58	-0.37079	0.13026	-0.35554	0.0062	0.1264	0.1108
	Age (category: 2 levels)	59	-0.16817	0.10888	-0.20044	0.1280	0.0402	0.0233
	Diabetes duration (year)	59	-0.01659	0.01658	-0.13132	0.3215	0.0172	0.0000

When the dependent variable was log (GDR divided by free insulin), it significantly correlated with DBP, LDL, BMI, waist circumference, waist/height ratio, waist percentile, adiponectin, leptin, adiponectin/leptin ratio, gender, waist circumference (7 categories), waist circumference (2 categories) and central obesity. Among these predictors, waist/height ratio ($r=0.286$, $p = < 0.0001$) had the highest correlation with log (GDR divided by free insulin). In addition, central obesity ($r=0.268$, $p = < 0.0001$) and waist circumference (2 categories) ($r=0.126$, $p = 0.0062$) also highly correlated with log (GDR divided by free insulin).

8.0 MULTIPLE REGRESSION MODEL

8.1.1 Possible methods

There are different methods that can be used to improve the prediction performance of regression models, such as Least absolute shrinkage and selection operator (LASSO), Principal component analysis (PCA), and Cross-Validation (CV). We select different methods for different purposes.

Least absolute shrinkage and selection operator (LASSO) is a shrinkage and selection method for linear regression. It minimizes the residual sum of square to the sum of the absolute value of the coefficients being less than a constant, and it tends to produce some coefficients that are equal to zero. It has many advantages. For example, it is a good way to deal with collinearity problem. Because we had about 20 potential predictors, it is easy to cause collinearity issue. In addition, it also can find the most valuable predictors.

Principal component analysis (PCA) is a variable reduction procedure. Linear combination of the original predictors is used to create a new variable. It is a good way to develop a small number of artificial variables based on a large number of potential predictors, and these combined predictors can account for most of the variance among all the variables. We had about 20 potential predictors, and only had 60 observations. Fewer predictors can reduce the

risk of causing overfitting problem, so PCA was a potential method that could be incorporated into our analysis.

For our project, we wanted to come up a model that was easy to interpret in a clinical setting by medical personnel, and could predict the insulin sensitivity as well. Based on this purpose, Lasso and PCA were not ideal methods.

8.1.2 Cross-Validation

Regression with stepwise or backward selection method is a potential method that can be used to select predictors and build models. This method uses entire dataset, and the performance is evaluated using the same dataset, so the results of testing performance may not be correct, and this means if we had new data, the model cannot predict well. Stepwise selection starts with no predictor, and then enters and removes predictors, in a stepwise manner, until there is no variable to enter or remove anymore based on P-values. Backward selection begins with all the predictors, and then removes predictors until no statistically insignificant variable is included.

The situation when it is not possible to predict performance of the model by using yet-unseen data is called overfitting. We wanted to split the whole dataset into two parts, and one to build the model, and another one to test the performance. If we split the dataset into two equal parts, it is not practical to conduct analysis, because we have to build the model using only 30 observations. Even though this partition is not ideal, the idea behind this method is good to our analysis. This idea forms the basis of the Cross -Validation.

For the Cross - validation method, the dataset was split in to different functional sets, training set, validation set and test set. For our project, we just split the dataset into training set and test set. Training set is the dataset where models are fit on. The test sets are used to predict

the performance of the dataset and its generalizability. It is better than the regression methods based on all the dataset, because it can deal with the overfitting problem, and predict the performance of models.

8.1.2.1 Introduction

There are different types of cross-validation, such as k-fold CV, Leave-one-out CV (LOOCV). In k-fold CV, the traditional values of k are 5 and 10. The data is partitioned into k part, the model is built based on k-1 partitions of the data, and the performance of the selected model is tested on the remaining part of the data. The performance measure reported by k-fold CV is the average of the k iterations of the procedure. For LOOCV, n-1 points are data is used to build the model, and the remaining one data will be used to test the performance. The performance of the selected model is the average of n iterations of the procedure. As a general rule, empirical evidence suggested 5- and 10-fold CV would be preferred to LOOCV, but there is no consistent agreement on which types of Cross - Validation will be the best to perform analysis. In this project, we tried 5- fold, 10- fold, and Leave-one-out CV to build the models.

8.1.2.2 Criteria

Criteria used to select the model and stop the selection process are crucial issues for CV. Potential criteria are as following: ADJRSQ (adjusted R-square statistics), AIC (Akaike information criterion), AICC (corrected AIC), BIC (Sawa Bayesian information criterion), CP (Mallows Cp statistics), PRESS (predicted residual sum of square statistic), SBS (Schwarz Bayesian information criterion), SL (significance level of the F statistics), VALIDATE (average square error over the validation data), etc. These criteria can be used to specified STOP, CHOOSE, and SELECT procedure. The STOP procedure indicates the procedure should stop at

which step. The CHOOSE procedure decides how to choose the predictors. The SELECT procedure shows how to select the best model. In this project, we used CV PRESS to select the best model, and SL to select predictors. Because the CV PRESS indicates the prediction error, it would be better to be the selection criterion.

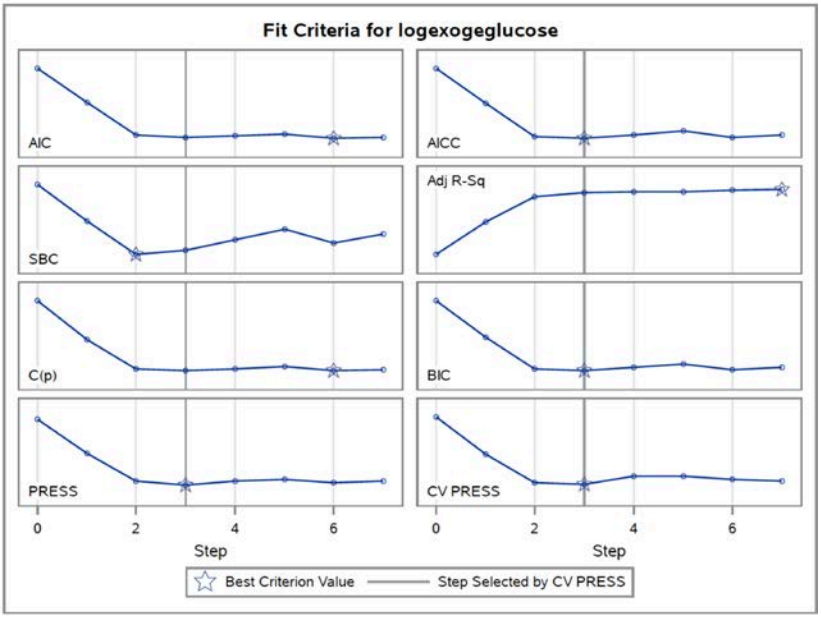
8.1.2.3 Example

In order to test CV procedure in SAS works, an example is given.

Table 10 Example

Regression with stepwise selection	Predictor	Coefficient		Regression with stepwise selection, 10-fold CV	Predictor	Coefficient
	Age	-0.026			BMI(%)	-0.007
	DBP(%)	-0.007			DBP(%)	-0.007
	Cholesterol	-0.003			Triglyceride	-0.002
	HDL	0.008			Intercept	3.099
	BMI(%)	-0.005				
	Intercept	3.219				

In this illustrative example, I used the stepwise selection method (p-enter = 0.35; p-stay = 0.35). The first result was from regression without using CV, and the second result was from regression with using CV. CV choose different predictors in the model compared to the method without CV.



Note: GDR=exogenous-glucose

Figure 2 Selection criteria

Because the criterion we chose was based on Cross-Validation predicted residual sum of square statistic (CV PRESS), from the figure above, we can see the best model did not come from the last step, on the other hand, it stopped at the step 3, which had the smallest CV PRESS. If we select the adjusted R-square as the criterion, the procedure will stop at the step 7, and the best model will come from the step 7.

Examining the figure above, we see that the predicted residual sum of square statistics does not improve after step 3 when using CV. In contrast without CV process would have continued to step 7 and selected a different set of variables.

9.0 RESULTS

Separate sets of models were separately built based on two dependent variables, Log (GDR) and Log (GDR divided by free insulin). The models not including the research variables are called practical models, and those including the research variables are called research models. Models were built using different methods, including regression models with stepwise selection method, regression models with backward selection method, regression models with stepwise selection method using 5- or 10- Cross-Validation, and regression models with stepwise selection method using leave-one-out Cross-Validation method. As mentioned previously, since waist, BMI percentile, waist/height ratio and waist percentile were highly correlated with each other, they were not included simultaneously in the same models separately. In Cross-Validation procedure, we chose models based on Cross-Validation predicted residual sum of squares statistic, and used significant level (P-enter = 0.15, P-stay=0.1) to select variables. The practical model was a model without considering leptin, adiponectin and adiponectin/leptin ratio. The research model was a model with considering leptin, adiponectin and adiponectin/leptin ratio, whereas the research considered these variables.

9.1.1 Log (GDR)

We built the models using methods mentioned above to predict log (GDR). Then we selected the best model, checked the assumptions, and conducted the diagnostic procedures.

9.1.1.1 Best model

Table 11 The best practical model for log (GDR)

Predictor	DBP percentile	SBP percentile	Gender	Diabetes duration	Waist circumference	Intercept
Coefficient	-0.005	-0.002	-0.238	-0.016	-0.020	4.421
P-value	0.016	0.142	0.0004	0.106	< .0001	< .0001

After comparing the adjusted R-square, AIC, and BIC, we chose the best practical model that included DBP percentile, SBP percentile, gender, diabetes duration, and waist circumference (R-square = 0.586, Adjusted R-square = 0.547, AIC = -103.359, BIC = -160.437). The directions of the relationship between each predictor and log (GDR) made sense from a clinical perspective.

Table 12 The best research model for log (GDR divided by free insulin)

Predictor	DBP percentile	Gender	Diabetes duration	Waist circumference	Adiponectin /leptin	Intercept
Coefficient	-0.006	-0.149	-0.017	-0.015	0.042	3.794
P-value	0.001	0.042	0.091	< .0001	0.0385	< .0001

In contrast, the variables in the best research model included DBP percentile, gender, diabetes duration, waist circumference, and adiponectin/leptin ratio (R-square = 0.602, Adjusted R-square = 0.565, AIC = -105.749, BIC = -162.472). The directions of each predictor and log (GDR) had reasonable clinical meanings.

9.1.1.2 Check assumptions

The assumptions for linear regression models are:

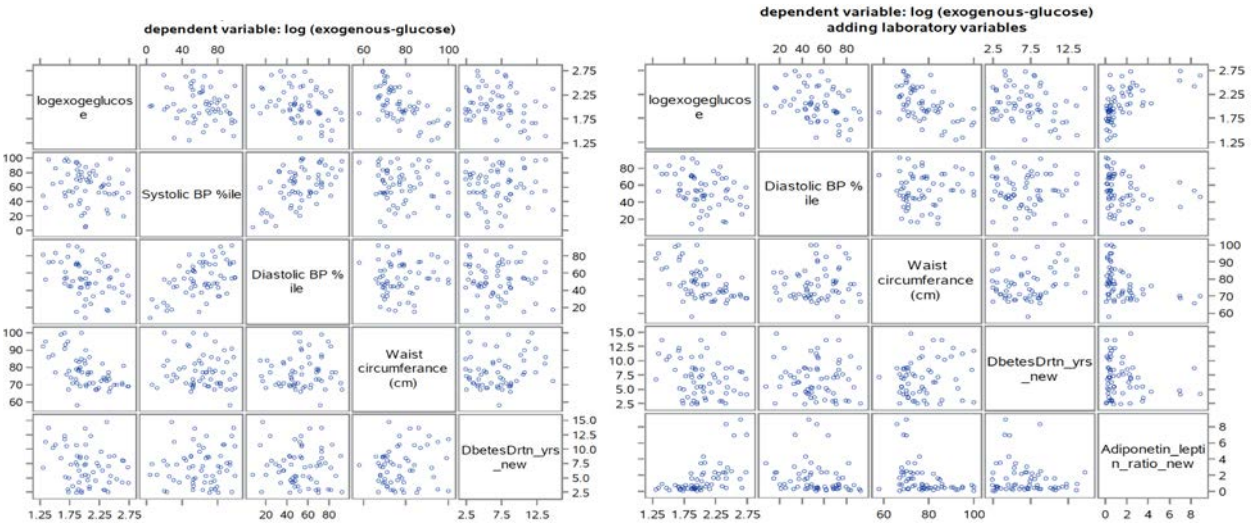
1. Existence: For each specific combination of the fixed x's, y is a random variable with a certain probably distribution.
2. Independent: The y values are statistically independent of each other.
3. Linearity: The mean of y for each specific combination of x1, x2, x3, … , xk is a linear function of x1, x2, x3, … , xk.
4. Homoscedasticity: The variance of y is the same for any fixed combination of x1, x2, x3, x4, … , xk.
5. Normality: For any fixed combination of x1, x2, x3, … , xk, the random variable y has a normal distribution. Residuals should be normal distributed.
6. Collinearity: There should not be collinearity problem.

- **Existence & Independent**

According to the study design, no more than one children came form the same family, and there was no relationship among children, so all the eligible children are independent of each other. In addition, we assumed dependent variables, GDR and GDR divided by free insulin, had their own distributions, so the assumption of existence is also met.

- **Linearity**

Plots were drawn of each predictor with log (GDR) in order to help researchers investigate their relationships.



Note: GDR=exogenous-glucose

Figure 3 Scatterplot matrix for predictors and log (GDR)

The plots were fine to meet the linearity assumption, and there is no obvious evidence to violate it. Even though they were not perfect linear relationships, this was probably because of small sample size.

- **Homoscedasticity**

Homoscedasticity means the variances of the residuals are consistent.

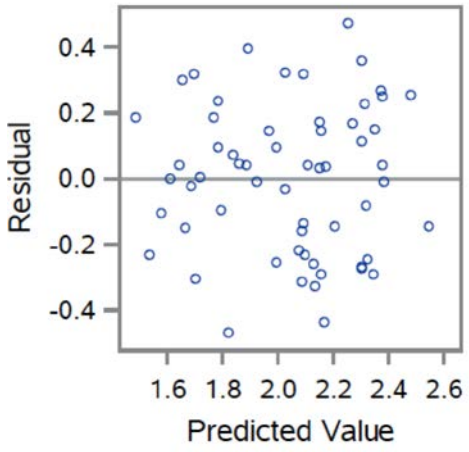


Figure 4 Residuals distribution for practical model

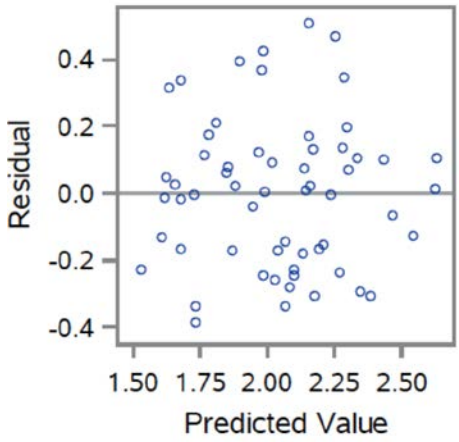


Figure 5 Residuals distribution for research model

Because there is no pattern in these two residuals graphs, these two models met the assumption of homoscedasticity.

- **Normality**

The residuals should be normally distributed. I used the QQ-plot to investigate the normality assumption.

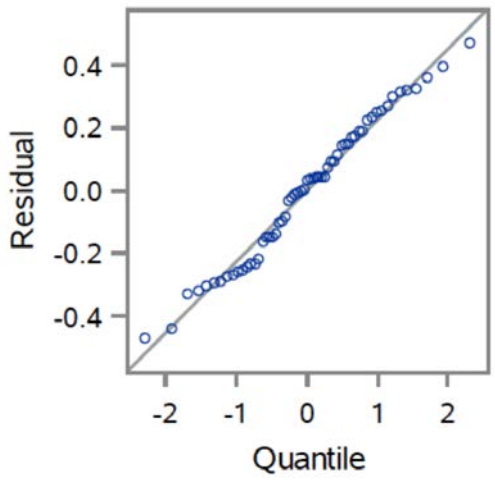


Figure 6 QQ-Plot for practical model

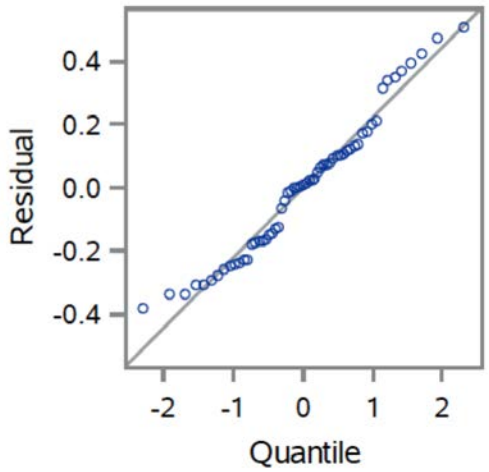


Figure 7 QQ-Plot for research model

The distribution of residuals indicated that residuals were normally distributed, so this assumption was validated.

- **Collinearity**

Collinearity means there is a linear relationship among the predictors, and two variables are linear combinations of one another predictor. Collinearity can cause unstable coefficients and inflated standard errors. Collinearity checking was performed using variance inflation factors (VIF). If VIF is greater than 10, there is a severe collinearity problem. Otherwise, no severe collinearity problem is presented.

Table 13 VIF for practical model

Predictor	DBP percentile	SBP percentile	Gender	Diabetes duration	Waist circumference
VIF	1.393	1.386	1.046	1.093	1.091

Table 14 VIF for research model

Predictor	DBP percentile	Gender	Diabetes duration	Waist circumference	Adiponectin/leptin ratio
VIF	1.065	1.379	1.074	1.245	1.674

Because all the VIFs were less than 10, there was no collinearity problem in these two models.

9.1.1.3 Diagnostics

The criteria below were the cut-points used to perform individual diagnostics on individual data points. K is the number of predictors, and n is number of observations.

Outliers are observations with large residuals, and we can be tested using the Studentized residuals (RStudent).

High leverage is an observation with an extreme value for a predictor variable. Leverage measures how far an observation deviates from the mean of that variable, and it can be evaluated through the leverage cut-point, (2k+2)/n.

Influential point is called when the observation substantially change the estimates of coefficient, and it can be checked using the Cook's D, 4/n.

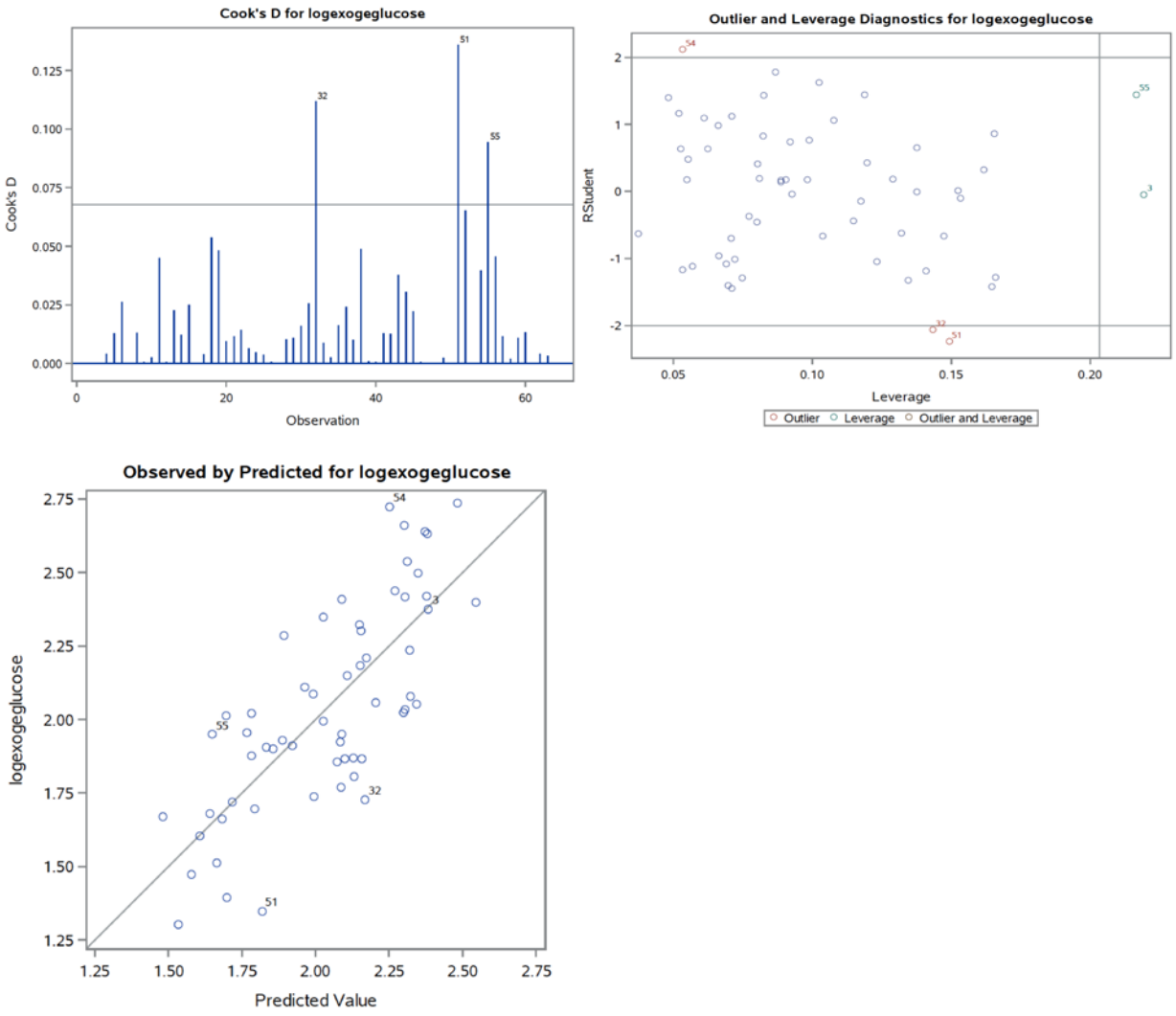
DFFITS measures the degree of an observation affect its fitted values, and it can be investigated through the cut-point, 2*sqrt(k/n), of DFFITS.

- **Practical model with dependent variable was log (GDR)**

Table 15 Criteria of diagnostics for practical model

Measure	Value	Cut-point
Leverage	> (2k+2)/n	> 0.237
Abs (Studentized residuals)	> 2	> 2
Cook's D	> 4/n	> 0.068

Note: k = 6, and n = 59.



Note: GDR=exogenous-glucose

Figure 8 Graphs for diagnostics

The plot of log (GDR) with predicted values indicated that our model fit well, because almost all the points were laid near the 45-degree line. The Cook's D graph showed patients 32, 51 and 55 might be considered as influential points. From the graph that includes information of student residuals, even though patients 32, 51 and 54 were above the cut-point, they were not far from that cut-point line, so these three observations were still considered reasonable, and would

not be considered outliers. Patients 3 and 55 seemed to be high leverage points, but because they were close to the cut-point line, it did not need to be considered as high leverage points.

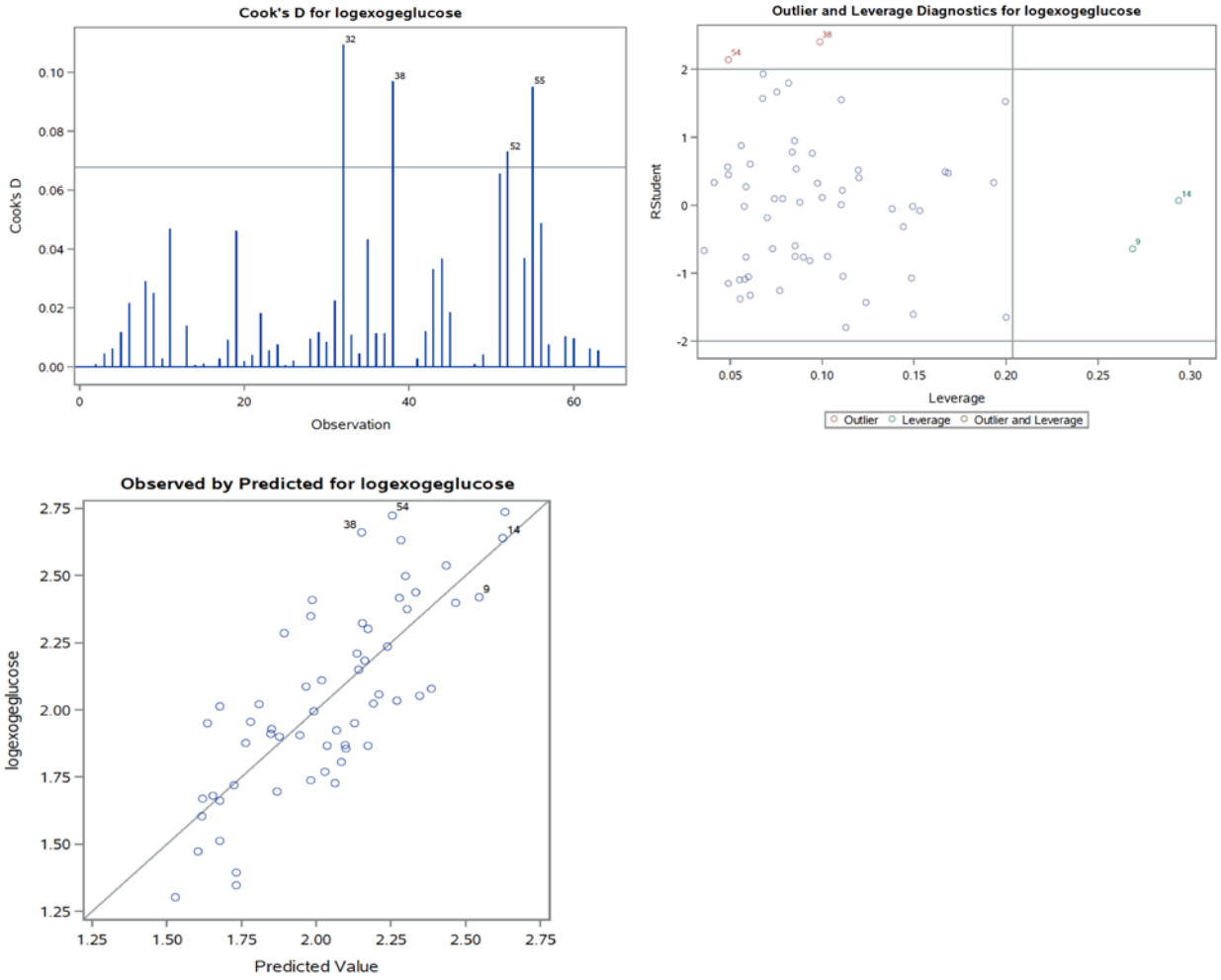
Research model with dependent variable log (GDR)

- **Research model with dependent variable was log (GDR)**

Table 16 Criteria of diagnostics for research model

Measure	Value	Cut-point
Leverage	> (2k+2)/n	> 0.237
Abs (Studentized residuals)	> 2	> 2
Cook's D	> 4/n	> 0.068

Note: k equaled to 6, and n was 59.



Note: GDR=exogenous-glucose

Figure 9 Graphs of diagnostics for research model

The results of the Cook's D graph indicated that patients 32, 38, and 55 might be suspected as influential points. We did not consider patient 52 as an influential point, because it was not far from the cut-point line. The graph that describes student residuals and leverage points indicated that patients 9 and 14 should be treated as high leverage points. When exploring the outliers, even though patients 38 and 54 were above the cut-point line, because they were still near the cut-point line, we did not treat them as outliers.

9.1.2 Log (GDR divided by free insulin)

9.1.2.1 Best model

Table 17 The best practical model for log (GDR divided by free insulin)

Predictor	DBP percentile	Gender	Waist circumference	HbA1C at the time of study	Intercept
Coefficient	-0.006	-0.263	-0.022	-0.090	5.214
P-value	0.009	0.002	< .0001	0.052	< .0001

Similarly, we compared adjusted R-square, AIC, and BIC to choose the best practical and research models. The best practical model included DBP percentile, gender, HbA1C at the time of study, diabetes duration, and waist circumference (R-square = 0.516, Adjusted R-square = 0.48, AIC = -135.9, BIC = -133).

Table 18 The best research model for log (GDR divided by free insulin)

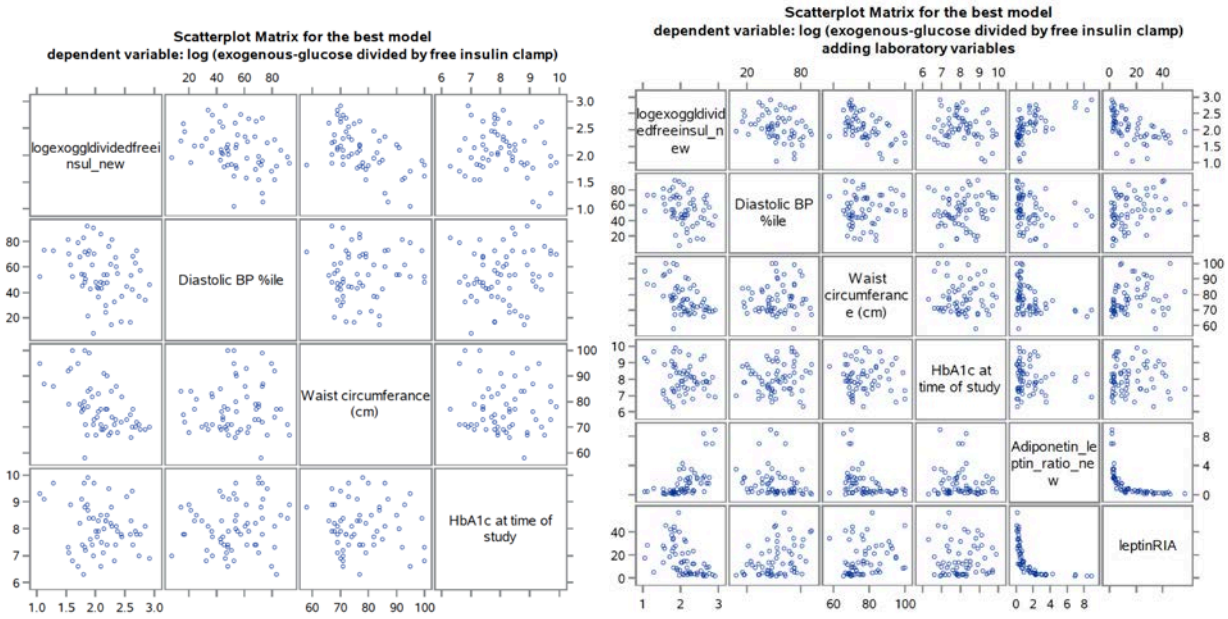
Predictor	DBP percentile	HbA1C at the time of study	Waist circumference	Adiponectin /leptin	Leptin	Intercept
Coefficient	-0.004	-0.084	-0.016	0.055	-0.007	4.249
P-value	0.082	0.055	0.0004	0.032	0.041	< .0001

The best research model included DBP percentile, HbA1C at the time of study, waist circumference, leptin, and adiponectin/leptin (R-square = 0.581, Adjusted R-square = 0.54, AIC = -82.174, BIC = -137.183).

9.1.2.2 Assumption

When the dependent variable was log (GDR divided by free insulin), the results of checking assumptions are presented below.

- **Linearity**



Note: GDR=exogenous-glucose

Figure 10 Scatterplot matrix for predictors and log (GDR divided by free insulin)

The above two graphs indicated there was no evidence to violate the assumption of linear relationship between predictors and dependent variables.

- **Homoscedasticity**

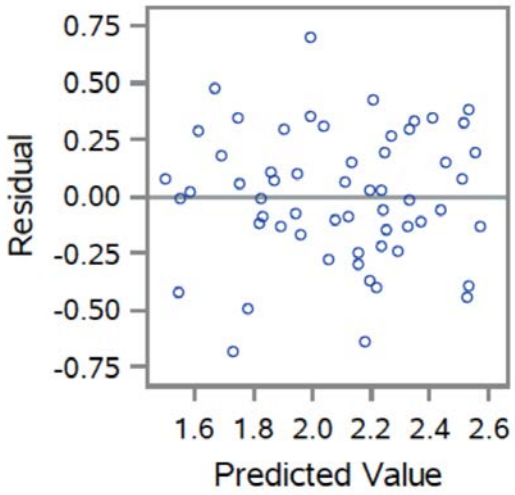


Figure 11 Residuals distribution for practical model

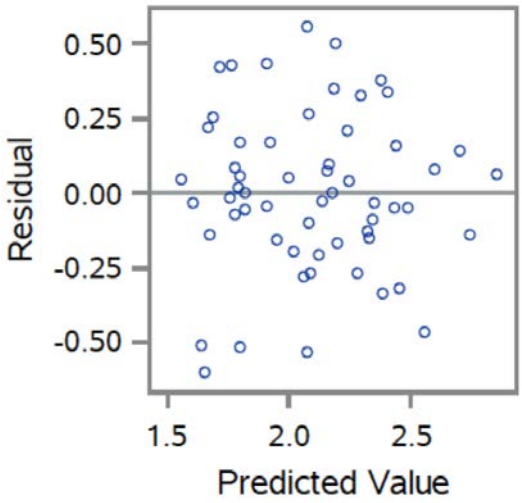


Figure 12 Residuals distribution for research model

All the residuals were random, and there is no pattern in the residuals graphs, so no model violated homoscedasticity assumption.

- **Normality**

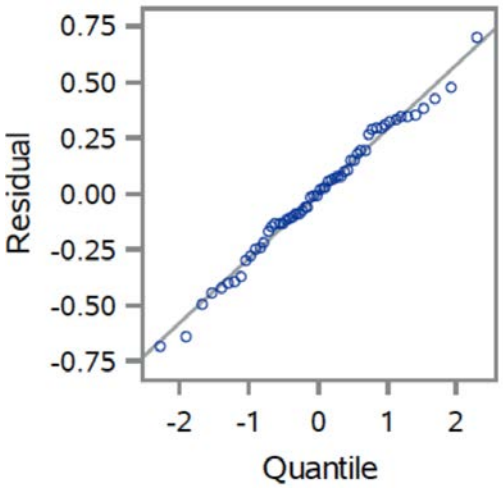


Figure 13 QQ-Plot for practical model

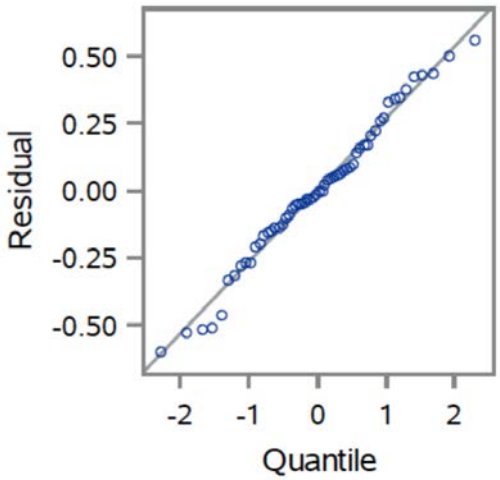


Figure 14 QQ-Plot for research model

Because all the points were on the 45-degree lines, so the results of the QQ-plots showed these two models met the normality assumption.

- **Collinearity**

Table 19 VIF for practical model

Predictor	DBP percentile	Gender	Waist circumference	HbA1C at the time of study
VIF	1.026	1.011	1.005	1.015

Table 20 VIF for research model

Predictor	DBP percentile	HbA1C at the time of study	Waist circumference	Adiponectin/leptin	Leptin
VIF	1.133	1.041	1.185	1.774	1.684

If the VIF was greater than 10, we should consider there was severe collinearity problem. Because all the VIFs were smaller than 10, collinearity problem did not present.

9.1.2.3 Diagnostics

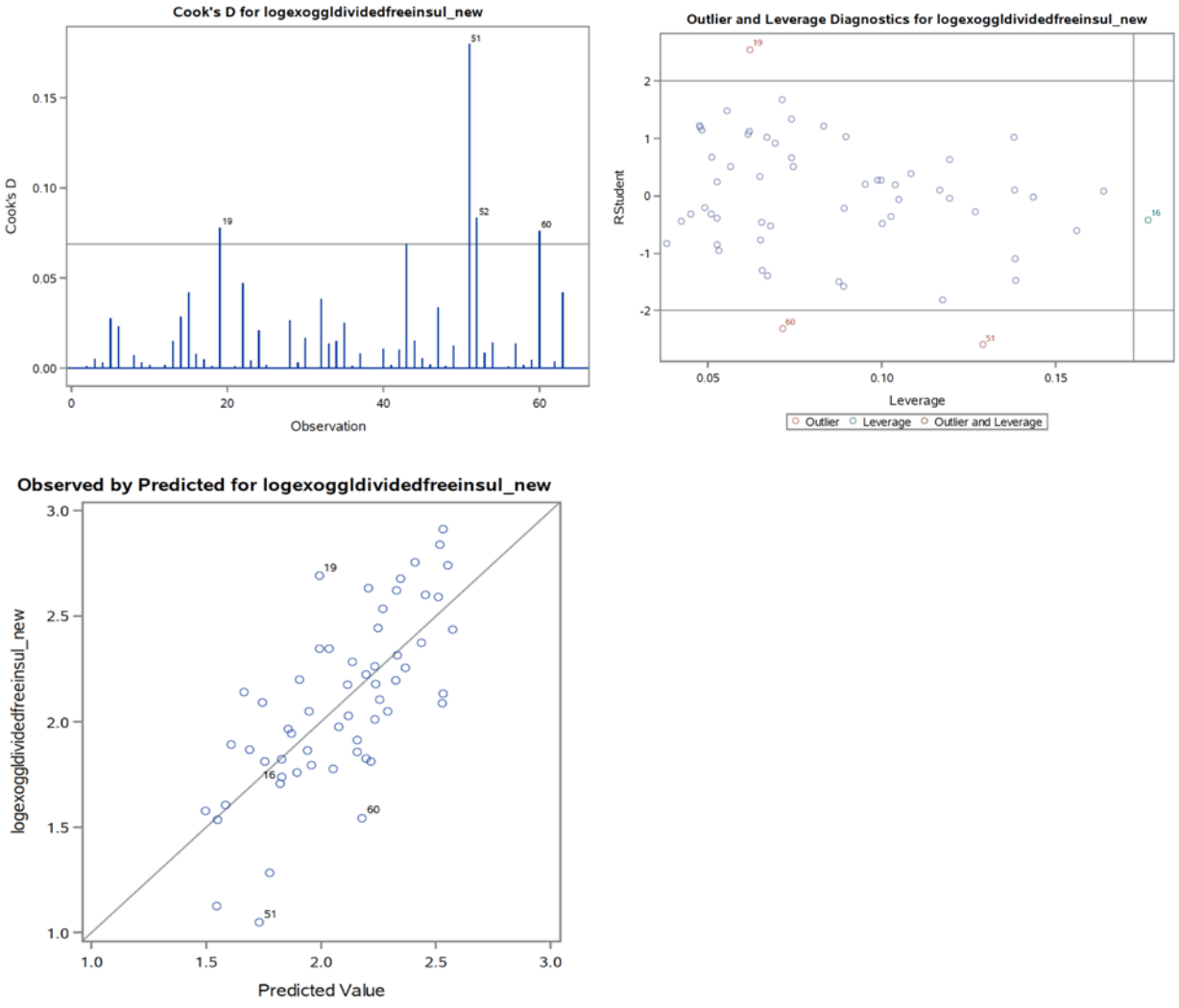
The criteria below were cut-points to investigate individual diagnostics. K was number of predictors, and n was number of observations.

- **Practical model with dependent variable was log (GDR divided by free insulin)**

Table 21 Criteria of diagnostics for practical model

Measure	Value	Cut-point
Leverage	> (2k+2)/n	> 0.207
Abs (Studentized residuals)	> 2	> 2
Cook's D	> 4/n	> 0.069

Note: k = 5, and n = 58.



Note: GDR=exogenous-glucose

Figure 15 Graphs of diagnostics for practical model

The results of the Cook's D indicated that the patient 51 could be an influential point. Because patients 19, 52, and 60 were not far from the cut-point line, we did not need to treat these observations as influential points. The above graph showed that patients 19, 51, and 60 may be outliers, but because they were not far from the cut-point lines, it was unnecessary to treat these points as outliers. The patient 16 might be considered as a high leverage point, but because the patient 16 was close to the cut-point line, we did not need to consider it as a high

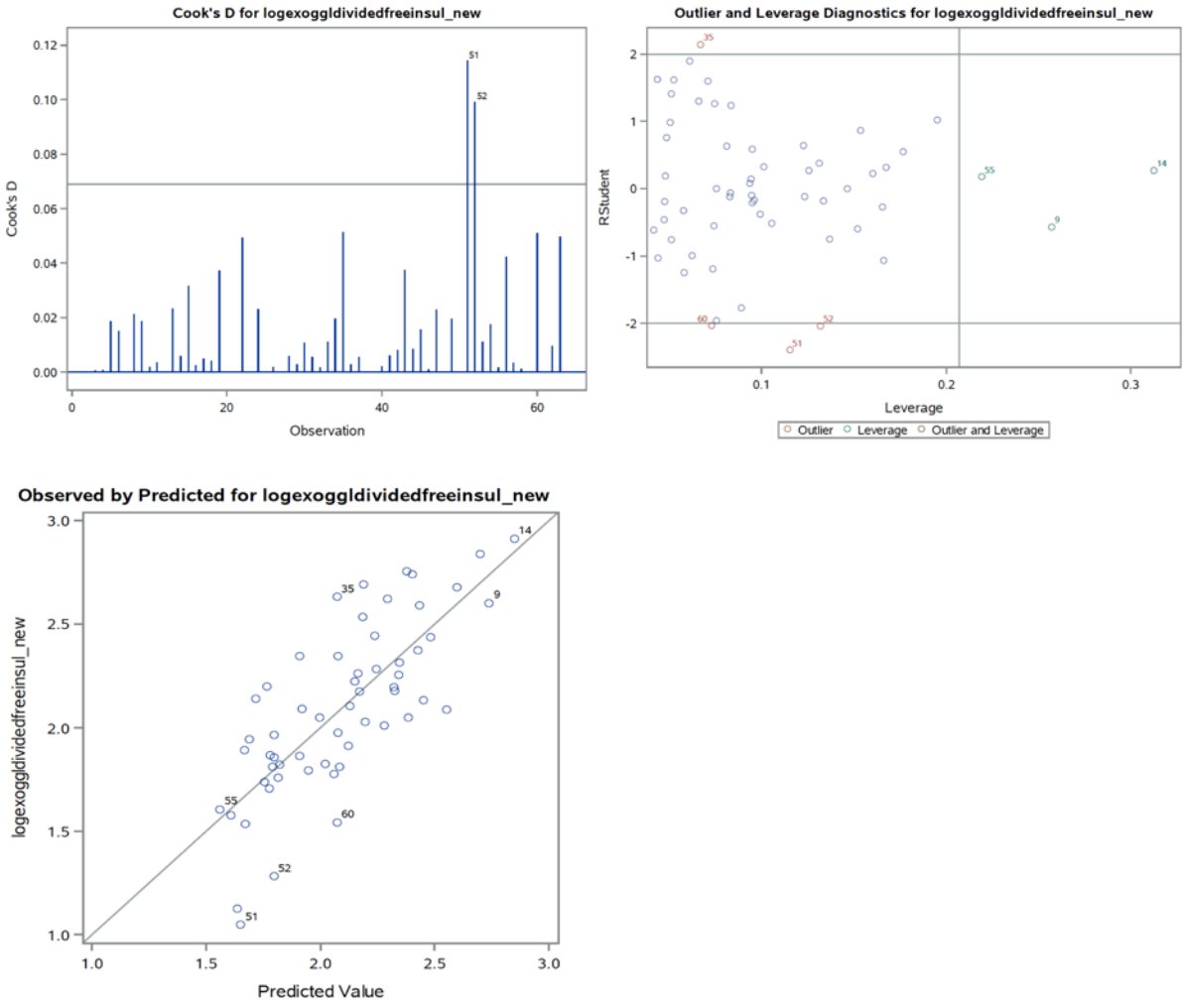
leverage point. From the plots of the log (GDR divided by free insulin) with predicted values, although some points were not near the 45-degree line, most of the points were close to the line. Therefore, it is reasonable to believe our model fit well.

- **Research model with dependent variable log (GDR divided by free insulin)**

Table 22 Criteria of diagnostics for practical model

Measure	Value	Cut-point
Leverage	> (2k+2)/n	> 0.241
Abs (Studentized residuals)	> 2	> 2
Cook's D	> 4/n	> 0.069

Note: k equaled to 6, and n was 58.



Note: GDR=exogenous-glucose

Figure 16 Graphs of diagnostics for research model

The Cook's D graph indicated patients 51 and 52 were influential points. When checking for outliers, although there were several points that were away from the cut-points, in general they were not that extreme, so we believed no outliers presented. The outlier and leverage graph indicated that the patient 14 should be a high leverage point, because it was far from the cut-point line. Overall, the plot of log (GDR divided by free insulin) with predicted values indicated that the model fitted well.

10.0 DISCUSSION AND CONCLUSION

Among all the 80 models we built, 72/80 (90%) models included DBP. The best model that had the highest R-square also included waist circumference. After just including DBP percentile and the waist circumference into the model, we found that it explained 42.61% of the variability in the log (GDR) (R-square = 0.4261, Adjusted R-square = 0.406). In addition, the DBP percentile by itself only explains 12.42% of the variability, and the waist circumference by itself explains 30.5% of the variability. When the dependent variable was log (GDR divided by free insulin), the model only including the DBP percentile and the waist circumference can explain 38.17% of variability in the log (GDR divided by free insulin) (R-square = 0.3817, Adjusted R-square = 0.3592). The model with DBP percentile only explained 9.6% of the variability in the log (GDR divided by free insulin), and the waist circumference by itself explained 28.47% of the variability.

The four best models identified meet the assumptions of the linear regression modeling. Log (GDR) resulted in and R-square for 0.602 for the practical model and 0.596 for the research model. As these values were essentially the same, we preferred to use the practical model. When the dependent variable was log (GDR divided by free insulin), the practical model had an R-squares of 0.516 and research model a value of 0.581 separately, so including the research variables can improve the predict performance by about seven percent. Thus the model with research predictors would be better.

When using log (GDR) as the dependent variable, the practical and research models had the common predictors of DBP percentile, gender, diabetes duration, and waist circumference. When using log (GDR divided by free insulin) as the dependent variable, the practical and research models also had the common predictors of DBP percentile, waist circumference, and HbA1C at the time of study.

There were several limitations to our study. First, our data set was relatively small with 60 children. Ideally, we would have liked to have a much larger data set. However, from a clinical perspective it is very difficult to recruit patients for this type of procedure and by research standards this was a substantial data set. In the future we hope to collaborate with other investigators and further evaluate and refine our models. Another limitation was the range in HbA1c values among the study participants. The investigators constrained entry into the study to participants with HbA1c less than 10% to avoid the effect of glucotoxicity. This may be the reason why HbA1c was not a statistically significant predictor.

As a comparison to other insulin sensitive prediction models found in the literature, we applied the SEARCH model to our data. Their model explained 30.4% of the variability when using their complete model and 30.2% when using their simplified model. Our proposed model explains much more of the variability in insulin sensitivity when using log (GDR).

The public health effect is identifying an IS predictive model based on routinely gathered clinical measurements and laboratory value is a valuable alternative to the invasive euglycaemic-hyperinsulinaemic clamp study. The current gold standard of insulin sensitivity, euglycaemic-hyperinsulinaemic clamp, is an invasive intravenous study requiring fasting overnight hospital study. The model makes it practical to use in epidemiological and screening studies.

APPENDIX: SAS PROGRAMMING

```
proc format ;
value  gender
1= 'male'
2= 'female' ;
value    lrecumbnt
             0  = "Standing height"
             1  = "Recumbent length";
value    lbivht
             0  = "Acceptable normal range"
             1  = "Too low"
             2  = "Too high";
value    lbivwt
             0  = "Acceptable normal range"
             1  = "Too low"
             2  = "Too high";
value    lbivwht
             0  = "Acceptable normal range"
             1  = "Too low"
             2  = "Too high";
value    lbivbmi
             0  = "Acceptable normal range"
             1  = "Too low"
             2  = "Too high";
value    lbivht_3months
             0  = "Acceptable normal range"
             1  = "Too low"
             2  = "Too high";
run ;
data dia;
set dia (drop=BMI BMIPCT BMIZ HAZ HCPCT HCZ HTPCT WAZ WHPCT WHZ WTPCT
case_ID dbp_score p_dbp_score p_sbp_score
sbp_score wc_perc);
run;
data dia (rename=(VAR00002=case_id));
set dia;
run;
proc sort data=dia;
by case_id;
proc sort data=tmp1.Anthro_bmi_waist_bp_scores_add;
by case_id;
run;
data final_dataset;
merge dia tmp1.Anthro_bmi_waist_bp_scores_add;
by case_id;
```

```
run;
data final_dataset_analysis;
set final_dataset (drop=bapwv birthweight build cfpwv crp
hcpct hcz hdltarget icam il6 imtingrid mfmdbts mfmlpd mfmobs mfmothr
mtrnldbts mtrnllpd mtrnlobs mtrnlothr pfamdbt pfamlpd pfamobs pfamother
ptrnldbts ptrnllpd ptrnlobs ptrnlothr ratioandroidtotalfat sat
sblingother
tnfalpha totalat vat vcam
_BIVBMI _BIVHC _BIVHT _BIVWHT _BIVWT percentbodyfatDEXA
percentbodyfatgynoid
percentfatandroid percentfatregionDEXA percentvatfat recumbnt whz whpct
);
run;
data final_dataset_analysis;
set final_dataset_analysis;
label adiponectinRIA_nov2013='adiponectinRIA_nov2013'
adiponectinRIA_oct2013='adiponectinRIA_oct2013';
run;
data final_dataset_analysis;
set final_dataset_analysis;
ageyrs_new=agemonths/12;
format ageyrs_new 5.2;
run;
data final_dataset_analysis;
set final_dataset_analysis;
if leptinRIA_oct2013^=. then leptinRIA=leptinRIA_oct2013;
else if leptinRIA_oct2013=. then leptinRIA=leptinRIA_aug2013;
run;
data final_dataset_analysis;
set final_dataset_analysis;
if adiponectinRIA_nov2013^=. Then
adiponectinRIA=adiponectinRIA_nov2013;
else if adiponectinRIA_nov2013=. then
adiponectinRIA=adiponectinRIA_oct2013;
run;
data final_dataset_analysis;
set final_dataset_analysis;
waist_height_ratio_new=waist/height;
tghdlratio_new=triglyceride/hdlt;
format tghdlratio_new 5.2;
if waist_height_ratio_new>0.5 then central_obesity_new=1;
else if .<waist_height_ratio_new<0.5 then central_obesity_new=0;
else central_obesity_new=.;
run;
data final_dataset_analysis;
set final_dataset_analysis;
exoggldividedfreeinsul_new=100*exogeglucose/freeinsulinclamp;
Adiponetin_leptin_ratio_new=adiponectinRIA/leptinRIA;
format exoggldividedfreeinsul_new 6.2;
run;
data final_dataset_analysis;
set final_dataset_analysis;
if wc_perc=0 then waist_percent=5;
else if wc_perc=1 then waist_percent=17.5;
else if wc_perc=2 then waist_percent=37.5;
else if wc_perc=3 then waist_percent=62.5;
else if wc_perc=4 then waist_percent=80;
```

```
else if wc_perc=5 then waist_percent=87.5;
else if wc_perc=6 then waist_percent=95;
run;
data final_dataset_analysis;
set final_dataset_analysis;
if .<wc_perc<5 then waist_categorical=1;
else if wc_perc>=5 then waist_categorical=2;
run;
data final_dataset_analysis;
set final_dataset_analysis;
if 11<ageyrs<15 and gender=2 then age_girl=1;
else if ageyrs>=15 and gender=2 then age_girl=2;
else age_girl=.;
run;
data final_dataset_analysis;
set final_dataset_analysis;
if 12<ageyrs<16 and gender=1 then age_boy=1;
else if ageyrs>=16 and gender=1 then age_boy=2;
else age_boy=.;
run;
data final_dataset_analysis;
set final_dataset_analysis;
if gender=1 then age_categorical=age_boy;
else if gender=2 then age_categorical=age_girl;
run;
data age_girl;
set final_dataset_analysis;
where gender=2;
run;
data age_boy;
set final_dataset_analysis;
where gender=1;
run;
proc means data=final_dataset_analysis n mean stderr median p5 p25  p75
p95 min max maxdec=3;
var logexogeglucose logexoggldividedfreeinsul_new;
class age_categorical;
run;

data final_dataset_analysis;
set final_dataset_analysis;
logexogeglucose=log(exogeglucose);
logexoggldividedfreeinsul_new=log(exoggldividedfreeinsul_new);
logexogglucosebyFFM=log(exogenousglucosebyFFM);
logexogglucosebyleanmass=log(exogenousglucosebyleanmass);
run;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\ttest_age_girl.pdf';
proc univariate data=age_girl;
var logexogeglucose;
qqplot;
run;
proc univariate data=age_boy;
var logexogeglucose;
qqplot;
run;
```

```
proc ttest data=age_boy;
var logexogeglucose;
class age_boy;
run;
proc ttest data=age_girl;
var logexogeglucose;
class age_girl;
run;
proc ttest data=final_dataset_analysis;
var logexogeglucose;
class age_categorical;
run;
proc ttest data=final_dataset_analysis;
var logexoggldividedfreeinsul_new;
class age_categorical;
run;
proc means data=age_boy n mean stderr median p5 p25  p75 p95 min max
maxdec=3;
class age_boy;
var logexogeglucose;
run;
proc means data=age_girl n mean stderr median p5 p25  p75 p95 min max
maxdec=3;
class age_girl;
var logexogeglucose;
run;
ods pdf close;

proc sql;
select age_categorical
from final_dataset_analysis;
quit;
data age_girl;
set final_dataset_analysis;
where gender=2;
run;
data age_boy;
set final_dataset_analysis;
where gender=1;
run;
data final_dataset_analysis;
set final_dataset_analysis;
tghdlratio_new=Triglyceride/hdlt;
adiponectin_leptin_ratio=adiponectinRIA/leptinRIA;
run;
ods pdf file='\\psf\Home\Documents\thesis\thesis_result\attribute.pdf';
proc contents data=final_dataset_analysis;
run;
ods pdf close;
data final_dataset_analysis;
set final_dataset_analysis;
DbetesDrtn_yrs_new=DbetesDrtn/12;
format DbetesDrtn_yrs_new 5.3;
run;
data final_dataset_analysis;
set final_dataset_analysis;
if gender=2 then tanner_stage=tanner_stage_breasts;
```

```
if gender=1 then tanner_stage=tanner_stage_pubic_hair;
run;
proc sql;
select tanner_stage, count(tanner_stage) as count,
tanner_stage_categorical
from final_dataset_analysis;
quit;
data final_dataset_analysis;
set final_dataset_analysis;
if tanner_stage<5 then tanner_stage_categorical=1;
else if tanner_stage=5 then tanner_stage_categorical=2;
run;
proc sql;
 select DbetesDrtn, count(DbetesDrtn) as count1,
  DbetesDrtn_yrs, count(DbetesDrtn_yrs) as count2,
  DbetesDrtn_yrs_new, count(DbetesDrtn_yrs_new) as count3
 from final_dataset_analysis;
quit;

ods pdf file='\\psf\Home\Documents\thesis\thesis_result\descriptive
statistics.pdf';
title 'Descriptive statistics of potential predictors';
proc means data=final_dataset_analysis n mean stderr median p5 p25  p75
p95 min max maxdec=3;
var Ageyrs DbetesDrtn_yrs_new bpdiastol p_dbp_score bpsystol
p_sbp_score Cholesterol hdlt ldl Triglyceride bmipct
waist_height_ratio_new waist_percent waist tghdlratio_new
adiponectin_leptin_ratio adiponectinRIA
leptinRIA Last_HbA1C Avrg_HbA1C;
label p_dbp_score='Diastolic BP(%)';
label p_sbp_score='Systolic BP(%)';
label waist_height_ratio_new='Waist/height';
label waist_percent='Waist(%)';
label DbetesDrtn_yrs_new='Duration';
where logexogeglucose^=.;
run;
proc means data=final_dataset_analysis n mean stderr median p5 p25  p75
p95 min max maxdec=3;
var Ageyrs DbetesDrtn_yrs_new bpdiastol p_dbp_score bpsystol
p_sbp_score Cholesterol hdlt ldl Triglyceride bmipct
waist_height_ratio_new waist_percent waist tghdlratio_new
adiponectin_leptin_ratio adiponectinRIA
leptinRIA Last_HbA1C Avrg_HbA1C;
label p_dbp_score='Diastolic BP(%)';
label p_sbp_score='Systolic BP(%)';
label waist_height_ratio_new='Waist/height';
label waist_percent='Waist(%)';
label DbetesDrtn_yrs_new='Duration';
where logexoggldividedfreeinsul_new^=.;
run;

proc means data=final_dataset_analysis n mean stderr median p5 p25  p75
p95 min max maxdec=3;
var exogeglucose exoggldividedfreeinsul_new;
run;
ods pdf close;
title;
```

```
proc format;
 value wc_perc_fmt 0='5%'
                   1='17.5%'
                           2='37.5%'
                           3='62.5%'
                           4='80%'
                           5='87.5%'
                           6='95%';
run;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\Frequency_categorical
variables.pdf';
title 'Frequency of categorical variables';
proc freq data=final_dataset_analysis;
table gender wc_perc central_obesity_new waist_categorical
tanner_stage_categorical;
format wc_perc wc_perc_fmt.;
where exogeglucose^=.;
run;
title;
ods pdf close;

proc template;
edit base.corr.stackedmatrix;
column (rowname rowlabel) (matrix) * (matrix2);
edit matrix;
cellstyle _val_  = -1.00 as {backgroundcolor=CXEEEEEE},
          _val_ <= -0.75 as {backgroundcolor=red},
          _val_ <= -0.50 as {backgroundcolor=yellow},
          _val_ <= -0.25 as {backgroundcolor=cyan},
          _val_ <=  0.25 as {backgroundcolor=white},
          _val_ <=  0.50 as {backgroundcolor=cyan},
          _val_ <=  0.75 as {backgroundcolor=yellow},
          _val_ <   1.00 as {backgroundcolor=red},
          _val_  =  1.00 as {backgroundcolor=CXEEEEEE};
end;
end;
run;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\correlation_predictors
and dependent variables.pdf';
title 'correlations between continuous vairables and logexogeglucose';
proc corr data=final_dataset_analysis;
 var Ageyrs DbetesDrtn_yrs_new p_dbp_score p_sbp_score Cholesterol hdlt
ldl Triglyceride bmipct
waist waist_height_ratio_new waist_percent tghdlratio_new
adiponectin_leptin_ratio adiponectinRIA
leptinRIA Last_HbA1C Avrg_HbA1C;
 with logexogeglucose;
label p_dbp_score='Diastolic BP(%)';
label p_sbp_score='Systolic BP(%)';
label waist_height_ratio_new='Waist/height';
label waist_percent='Waist(%)';
label DbetesDrtn_yrs_new='Duration(year)';
label Last_HbA1C='HbA1C at the time of study';
label Avrg_HbA1C='HbA1C at last 5 clinic visits';
```

```
label ageyrs='age(year)';
label bmipct='bmic(%)';
label adiponectinRIA='adiponectin';
label leptinRIA='leptin';
label tghdlratio_new='Triglyceride/HDL';
label adiponectin_leptin_ratio='adiponectin/leptin';
run;
title;
title 'correlations between continuous vairables and
logexoggldividedfreeinsul';
footnote 'using logexoggldividedfreeinsul_new';
proc corr data=final_dataset_analysis;
 var Ageyrs DbetesDrtn_yrs_new p_dbp_score p_sbp_score Cholesterol hdlt
ldl Triglyceride bmipct
waist waist_height_ratio_new waist_percent tghdlratio_new
adiponectin_leptin_ratio adiponectinRIA
leptinRIA Last_HbA1C Avrg_HbA1C;
 with logexoggldividedfreeinsul_new;
 label p_dbp_score='Diastolic BP(%)';
label p_sbp_score='Systolic BP(%)';
label waist_height_ratio_new='Waist/height';
label waist_percent='Waist(%)';
label DbetesDrtn_yrs_new='Duration(year)';
label Last_HbA1C='HbA1C at the time of study';
label Avrg_HbA1C='HbA1C at last 5 clinic visits';
label ageyrs='age(year)';
label bmipct='bmic(%)';
label adiponectinRIA='adiponectin';
label leptinRIA='leptin';
label tghdlratio_new='Triglyceride/HDL';
label adiponectin_leptin_ratio='adiponectin/leptin';
run;
title;
footnote;
title 'correlations between continuous vairables and
logexogglucosebyFFM';
proc corr data=final_dataset_analysis;
 var Ageyrs DbetesDrtn_yrs_new p_dbp_score p_sbp_score Cholesterol hdlt
ldl Triglyceride bmipct
waist waist_height_ratio_new waist_percent tghdlratio_new
adiponectin_leptin_ratio adiponectinRIA
leptinRIA Last_HbA1C Avrg_HbA1C;
 with logexogglucosebyFFM;
run;
title;
title 'correlations between continuous vairables and
logexogglucosebyleanmass';
proc corr data=final_dataset_analysis;
 var Ageyrs DbetesDrtn_yrs_new p_dbp_score p_sbp_score Cholesterol hdlt
ldl Triglyceride bmipct
waist waist_height_ratio_new waist_percent tghdlratio_new
adiponectin_leptin_ratio adiponectinRIA
leptinRIA Last_HbA1C Avrg_HbA1C;
 with logexogglucosebyleanmass;
run;
title;
ods pdf close;
```

```
proc template;
delete base.corr.stackedmatrix;
run;

data corrlation;
set final_dataset_analysis (keep= Ageyrs DbetesDrtn_yrs_new p_dbp_score
p_sbp_score Cholesterol hdlt ldl Triglyceride bmipct
waist waist_height_ratio_new waist_percent tghdlratio_new
adiponectin_leptin_ratio adiponectinRIA
leptinRIA Last_HbA1C Avrg_HbA1C);
run;

proc template;
edit base.corr.stackedmatrix;
column (rowname rowlabel) (matrix) * (matrix2);
edit matrix;
cellstyle _val_  = -1.00 as {backgroundcolor=CXEEEEEE},
          _val_ <= -0.75 as {backgroundcolor=red},
          _val_ <= -0.50 as {backgroundcolor=blue},
          _val_ <= -0.25 as {backgroundcolor=cyan},
          _val_ <=  0.25 as {backgroundcolor=white},
          _val_ <=  0.50 as {backgroundcolor=cyan},
          _val_ <=  0.75 as {backgroundcolor=blue},
          _val_ <   1.00 as {backgroundcolor=red},
          _val_  =  1.00 as {backgroundcolor=CXEEEEEE};
end;
end;
run;4126211134          347 south bouquet
ods pdf
body='\\psf\Home\Documents\thesis\thesis_result\correlation_matrix.pdf'
style=statistical;
title 'Correlation Matrix';
proc corr data=corrlation;
 ods select pearsoncorr;
 run;
 title;
 ods pdf close;
 proc template;
 delete base.corr.stackedmatrix;
 run;

data corrlation_suspected;
set final_dataset_analysis (keep=bmipct waist waist_height_ratio_new
waist_percent);
run;

proc template;
edit base.corr.stackedmatrix;
column (rowname rowlabel) (matrix) * (matrix2);
edit matrix;
cellstyle _val_  = -1.00 as {backgroundcolor=CXEEEEEE},
          _val_ <= -0.75 as {backgroundcolor=red},
          _val_ <= -0.50 as {backgroundcolor=blue},
          _val_ <= -0.25 as {backgroundcolor=cyan},
          _val_ <=  0.25 as {backgroundcolor=white},
          _val_ <=  0.50 as {backgroundcolor=cyan},
          _val_ <=  0.75 as {backgroundcolor=blue},
```

```
          _val_ <   1.00 as {backgroundcolor=red},
          _val_  =  1.00 as {backgroundcolor=CXEEEEEE};
end;
end;
run;
ods pdf
body='\\psf\Home\Documents\thesis\thesis_result\correlation
suspected_matrix.pdf'
style=statistical;
title 'Correlation Matrix';
proc corr data=corrlation_suspected;
 ods select pearsoncorr;
 run;
 title;
 ods pdf close;
 proc template;
 delete base.corr.stackedmatrix;
 run;

*using logexogelucose, logexoggldividedfreeinsul_new;
ods pdf
file="\\psf\Home\Documents\thesis\thesis_result\Univariate regression
models_&var1.pdf";
%let var1=logexogglucosebyleanmass;
proc reg data=final_dataset_analysis;
model &var1=ageyrs/stb;
plot &var1*ageyrs;
run;
proc reg data=final_dataset_analysis;
model &var1=p_dbp_score/stb;
plot &var1*p_dbp_score;
run;
proc reg data=final_dataset_analysis;
model &var1=p_sbp_score/stb;
plot &var1*p_sbp_score;
run;
proc reg data=final_dataset_analysis;
model &var1=Cholesterol/stb;
plot &var1*Cholesterol;
run;
proc reg data=final_dataset_analysis;
model &var1=hdlt/stb;
plot &var1*hdlt;
run;
proc reg data=final_dataset_analysis;
model &var1=ldl/stb;
plot &var1*ldl;
run;
proc reg data=final_dataset_analysis;
model &var1=Triglyceride/stb;
plot &var1*Triglyceride;
run;
proc reg data=final_dataset_analysis;
model &var1=bmipct/stb;
plot &var1*bmipct;
run;
proc reg data=final_dataset_analysis;
```

```
model &var1=waist/stb;
plot &var1*waist;
run;
proc reg data=final_dataset_analysis;
model &var1=waist_height_ratio_new/stb;
plot &var1*waist_height_ratio_new;
run;
proc reg data=final_dataset_analysis;
model &var1=waist_percent/stb;
plot &var1*waist_percent;
run;
proc reg data=final_dataset_analysis;
model &var1=tghdlratio_new /stb;
plot &var1*tghdlratio_new ;
run;
proc reg data=final_dataset_analysis;
model &var1=adiponectin_leptin_ratio /stb;
plot &var1*adiponectin_leptin_ratio ;
run;
proc reg data=final_dataset_analysis;
model &var1=adiponectinRIA /stb;
plot &var1*adiponectinRIA ;
run;
proc reg data=final_dataset_analysis;
model &var1=leptinRIA/stb;
plot &var1*leptinRIA;
run;
proc reg data=final_dataset_analysis;
model &var1=Last_HbA1C/stb;
plot &var1*Last_HbA1C ;
run;
proc reg data=final_dataset_analysis;
model &var1=Avrg_HbA1C/stb;
plot &var1*Avrg_HbA1C;
run;
proc reg data=final_dataset_analysis;
model &var1=gender/stb;
plot &var1*gender ;
run;
proc reg data=final_dataset_analysis;
model &var1=wc_perc /stb;
plot &var1*wc_perc;
run;
proc reg data=final_dataset_analysis;
model &var1=central_obesity_new /stb;
plot &var1*central_obesity_new ;
run;
proc reg data=final_dataset_analysis;
model &var1=waist_categorical/stb;
plot &var1*waist_categorical;
run;
proc reg data=final_dataset_analysis;
model &var1=age_categorical/stb;
plot &var1*age_categorical;
run;
proc reg data=final_dataset_analysis;
model &var1=DbetesDrtn_yrs_new/stb;
```

```
plot &var1*DbetesDrtn_yrs_new;
run;
ods pdf close;

*********Practical model using log (exogenous-glucose) as the dependent
variable**********************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_bmi
pct_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.1
and p_remove=.08*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.1
and p_remove=.08*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
```

```
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(60)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;
**********************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.1
and p_remove=.08*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.1
and p_remove=.08*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=backward
sls=0.1;
run;
```

```
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*****************************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_height_simple.pdf';
ods graphics on;
```

```
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
```

```
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*****************************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_cen
tral_obesity_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
```

```
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*******************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_percent_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
```

```
waist_percent/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
```

```
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_2_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
```

```
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_7_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;
```

```
ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender
DbetesDrtn_yrs_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender
DbetesDrtn_yrs_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender
```

```
DbetesDrtn_yrs_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(60)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

****Research model using log (exogenous-glucose) as the dependent
variable*****;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_bmi
pct_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.1
and p_remove=.08*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv
select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
```

```
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv
select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv
select=sl sle=.15 sls=.1)
cvMethod=split(60)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;
*********************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
```

```
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv
select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv
select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv
select=sl sle=.15 sls=.1)
cvMethod=split(59)
```

```
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*****************************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_height_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using stepwise selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
```

```
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

****************************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_cen
tral_obesity_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new central_obesity
leptinRIA adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;
```

```
ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new central_obesity
leptinRIA adiponectinRIA Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new central_obesity
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new central_obesity
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
```

```
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new central_obesity
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*****************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_percent_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist_percent
leptinRIA adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist_percent
leptinRIA adiponectinRIA Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist_percent
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(5)
stats=all
```

```
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist_percent
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist_percent
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_2_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA
Adiponetin_leptin_ratio_new waist_categorical/selection=stepwise
slstay=0.1 slentry=0.15 ;
```

```
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA
Adiponetin_leptin_ratio_new waist_categorical/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA
Adiponetin_leptin_ratio_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
```

```
ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(60)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;


ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexglucose_wai
st_7_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
```

```
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexogeglucose=Ageyrs p_dbp_score p_sbp_score hdlt ldl
tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(60)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;


*********Practical model using log (exogenous-glucose divided by free
insulin) as the dependent variable*****;
```

```
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivinsu
l_bmipct_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=stepwise slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
```

```
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
bmipct/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;
*********************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
```

```
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(58)
stats=all
orderselect
cvdetails=cvpress;
title;
ods graphics off;
ods pdf close;

*****************************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_height_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=stepwise
```

```
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
```

```
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(58)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*****************************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_central_obesity_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
```

```
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(58)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*******************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_percent_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
```

```
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent/selection=stepwise(choose=cv select=sl sle=.15 sls=.1)
cvMethod=split(58)
stats=all
orderselect
cvdetails=cvpress;
```

```
run;
title;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_2_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
```

```
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_7_simple.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender
DbetesDrtn_yrs_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
```

```
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender
DbetesDrtn_yrs_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender
DbetesDrtn_yrs_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender
DbetesDrtn_yrs_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender
DbetesDrtn_yrs_new/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(59)
stats=all
orderselect
```

```
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;
*********Research model using log (exogenous-glucose divided by free
insulin) as the dependent variable*****;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivinsu
l_bmipct_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct
leptinRIA adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct
leptinRIA adiponectinRIA Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
```

```
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new bmipct
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;
*********************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
```

```
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist
leptinRIA adiponectinRIA Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new waist
leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
```

```
cvMethod=split(58)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*****************************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_height_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
```

```
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_height_ratio_new leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(58)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*****************************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_central_obesity_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
```

```
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
```

```
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
central_obesity leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(58)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

*****************************************;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_percent_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
```

```
waist_percent leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new
waist_percent leptinRIA adiponectinRIA
Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv select=sl
sle=.15 sls=.1)
cvMethod=split(58)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_2_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
```

```
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA
Adiponetin_leptin_ratio_new waist_categorical/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA
Adiponetin_leptin_ratio_new waist_categorical/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(10)
stats=all
orderselect
```

```
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class waist_categorical;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new
waist_categorical/selection=stepwise(choose=cv select=sl sle=.15
sls=.1)
cvMethod=split(58)
stats=all
orderselect
cvdetails=cvpress; *deal with collinearity? A: No;
run;
title;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\logexoggldivided
freeinsul_waist_7_lab.pdf';
ods graphics on;
/*linear regression model, using stepwise selection method, p_enter=.15
and p_remove=.1*/
title 'mutiple linear regression with stepwise selection method';
title2 'using p_enter=.15  p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise
slstay=0.1 slentry=0.15 ;
run;
title;
ods graphics off;

ods graphics on;
/*linear regression model, using backward selection method,
p_remove=.1*/
title 'mutiple linear regression with backward selection method';
title2 'with p_remove=.1';
proc reg data=final_dataset_analysis;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=backward
sls=0.1;
run;
title;
ods graphics off;

ods graphics on;
```

```
title 'mutiple linear regression with 5 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv
select=sl sle=.15 sls=.1)
cvMethod=split(5)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with 10 fold cross-validation';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv
select=sl sle=.15 sls=.1)
cvMethod=split(10)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;

ods graphics on;
title 'mutiple linear regression with LOOCV';
title2 'and stepwise selection method';
proc glmselect data=final_dataset_analysis
 plots(stepAxis=number)=(criterionPanel ASEPlot);
 class wc_perc;
 model logexoggldividedfreeinsul_new=Ageyrs p_dbp_score p_sbp_score
hdlt ldl tghdlratio_new Last_HbA1C gender DbetesDrtn_yrs_new leptinRIA
adiponectinRIA Adiponetin_leptin_ratio_new/selection=stepwise(choose=cv
select=sl sle=.15 sls=.1)
cvMethod=split(59)
stats=all
orderselect
cvdetails=cvpress;
run;
title;
ods graphics off;
ods pdf close;

proc reg data=final_dataset_analysis;
model logexogeglucose=p_dbp_score waist;
run;
```

```
proc reg data=final_dataset_analysis;
model logexoggldividedfreeinsul_new=p_dbp_score waist;
run;

**********Scatterplots, checking the linearity assumption**********;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\matrix1.pdf';
proc sgscatter data=final_dataset_analysis;
title 'Scatterplot Matrix for the best model';
title2 'dependent variable: log (exogenous-glucose)';
matrix logexogeglucose p_sbp_score p_dbp_score waist
DbetesDrtn_yrs_new;
run;
title;
ods pdf close;
quit;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\matrix2.pdf';
proc sgscatter data=final_dataset_analysis;
title 'Scatterplot Matrix for the best model';
title2 'dependent variable: log (exogenous-glucose divided by free
insulin clamp)';
matrix logexoggldividedfreeinsul_new p_dbp_score waist Last_HbA1C;
run;
title;
ods pdf close;
quit;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\matrix3.pdf';
proc sgscatter data=final_dataset_analysis;
title 'Scatterplot Matrix for the best model';
title2 'dependent variable: log (exogenous-glucose)';
title3 'adding laboratory variables';
matrix logexogeglucose p_dbp_score waist DbetesDrtn_yrs_new
Adiponetin_leptin_ratio_new;
run;
title;
ods pdf close;
quit;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\matrix4.pdf';
proc sgscatter data=final_dataset_analysis;
title 'Scatterplot Matrix for the best model';
title2 'dependent variable: log (exogenous-glucose divided by free
insulin clamp)';
title3 'adding laboratory variables';
matrix logexoggldividedfreeinsul_new p_dbp_score waist Last_HbA1C
Adiponetin_leptin_ratio_new
leptinRIA;
run;
title;
ods pdf close;
quit;
```

```
************check residuals, leverage, cook's d dffits********;
ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\disgnostics&assu
mption1.pdf';
ods graphics on;
proc reg data=final_dataset_analysis
plots=(diagnostics(stats=all) RStudentByLeverage(label)
                CooksD(label) Residuals(smooth)
                DFFITS(label) DFBETAS ObservedByPredicted(label));
model logexogeglucose=p_sbp_score p_dbp_score waist DbetesDrtn_yrs_new
gender/vif tol;
run;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\disgnostics&assu
mption2.pdf';
ods graphics on;
proc reg data=final_dataset_analysis
plots=(diagnostics(stats=all) RStudentByLeverage(label)
                CooksD(label) Residuals(smooth)
                DFFITS(label) DFBETAS ObservedByPredicted(label));
model logexoggldividedfreeinsul_new=p_dbp_score waist Last_HbA1C
gender/vif tol;
run;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\disgnostics&assu
mption3.pdf';
ods graphics on;
proc reg data=final_dataset_analysis
plots=(diagnostics(stats=all) RStudentByLeverage(label)
                CooksD(label) Residuals(smooth)
                DFFITS(label) DFBETAS ObservedByPredicted(label));
model logexogeglucose=p_dbp_score waist DbetesDrtn_yrs_new
Adiponetin_leptin_ratio_new gender/vif tol;
run;
ods graphics off;
ods pdf close;

ods pdf
file='\\psf\Home\Documents\thesis\thesis_result\models\disgnostics&assu
mption4.pdf';
ods graphics on;
proc reg data=final_dataset_analysis
plots=(diagnostics(stats=all) RStudentByLeverage(label)
                CooksD(label) Residuals(smooth)
                DFFITS(label) DFBETAS ObservedByPredicted(label));
model logexoggldividedfreeinsul_new=p_dbp_score waist Last_HbA1C
Adiponetin_leptin_ratio_new leptinRIA/vif tol;
run;
ods graphics off;
ods pdf close;
```

```
quit;

********Compare with the models in the SEARCH project*************;
data final_dataset_analysis;
set final_dataset_analysis;
insulin_sens_search_simple = 3.7339 - (0.02155 * waist);
insulin_sens_search = 4.64725 - (0.02032 * waist) -
(0.09779 * Avrg_HbA1C) - (0.00235 * triglyceride);
run;
proc corr data=final_dataset_analysis;
var insulin_sens_search_simple insulin_sens_search;
with logexogeglucose;
run;
proc corr data=final_dataset_analysis;
var insulin_sens_search_simple insulin_sens_search;
with logexoggldividedfreeinsul_new;
run;

ods pdf file='\\psf\Home\Documents\thesis
\thesis_result\models\search models.pdf';
proc reg data=final_dataset_analysis;
model logexogeglucose=waist;
run;
proc reg data=final_dataset_analysis;
model logexogeglucose=waist Avrg_HbA1C triglyceride;
run;
ods pdf close;
```

BIBLIOGRAPHY

1. "Your Guide to Diabetes: Type 1 and Type 2." National Diabetes Information Clearinghouse. National Institute of Digestive and Kidney Diseases, National Institutes of Health, 10 February. 2014. Web. 10 March. 2015. http://diabetes.niddk.nih.gov/dm/pubs/type1and2/

2. "Type 1 Diabetes: What is it?" KidsHealth. Nemours Foundation, April. 2012. Web. 10 March. 2015. http://kidshealth.org/parent/diabetes_center/diabetes_basics/type1.html#

3. "Type 1 diabetes in children." Riverside. Riverside Health System, 03 March. 2011. Web. 10 March. 2015. http://riversideonline.com/health_reference/teen-health/ds00931.cfm

4. Hess-Fischl, Amy. " What is insulin?" Endocrineweb. Vertical Health LLC, 15 March. 2015. Web. 20 March. 2015. http://www.endocrineweb.com/conditions/type-1-diabetes/what-insulin

5. "Nine Things that Improve Insulin Sensitivity: Accelerate Fat Loss & Build Muscle Faster!" Poliquin Group. Poliquin Inc, 17 December. 2014. Web. 10 March. 2015. http://www.poliquingroup.com/ArticlesMultimedia/Articles/Article/1267/Nine_Things_that_Improve_Insulin_Sensitivity_Accel.aspx

6. Bjornstad, Petter. "Insulin Sensitivity and Complications in Type 1 Diabetes: New Insights." WJD World Journal of Diabetes. 8.

7. Schnell, O., & Cappuccio, F. (2013). "Type 1 diabetes and cardiovascular disease." Cardiovascular Diabetology. doi:10.1186/1475-2840-12-156

8. Orchard, T. J.,T. Costacou, Kretowski, A., and R. W. Nesto, R. "Type 1 Diabetes and Coronary Artery Disease." Diabetes Care (2006), 2528-2538

9. Signgh, Bhawna. "Surrogate markers of insulin resistance: A review." WJD World Journal of Diabetes, 36-36.

10. Dabelea, D., R. B. D' Agostino, C. C. Mason, N. West, R. F. Hamman, E. J. Mayer-Davis, D. Maahs, G. Klingensmith, W. C. Knowler, and K.. Nadeau. "Development, validation and use of an insulin sensitivity score in youths with diabetes: the SEARCH for Diabetes in Youth study." Diabetologia (2010), 78-86.

11. "Frequently Asked Questions (FAQs) about BMI." American Heart Association. American Heart Association, 26 August. 2014. Web. 10 March. 2015.

http://www.heart.org/HEARTORG/GettingHealthy/WeightManagement/BodyMassIndex/Frequently-Asked-Questions-FAQs-about-BMI_UCM_307892_Article.jsp

12. “Waist to height ratio.” Health-calc Health & Fitness Calculators. http://www.health-calc.com/body-composition/waist-to-height-ratio

13. Vliet, Mariska Van, Josine C. Van Der Heyden, Michaela Diamant, Ines A. Von Rosenstiel, Roger K. Schindhelm, Martijn W. Heymans, Desiderius P. M. Brandjes, Jos H. Beijnen, Henk J. Aanstoot, and Henk J. Veeze. “Overweight children with type 1 diabetes have a more favourable lipid profile than overweight non-diabetic children.” European Journal of Pediatrics Eur J pediatr (2011): 493-498.

14. Saltiel, Alan R., and C. Ronald Kahn. “Insulin Signalling and the Regulation of Glucose and Lipid Metabolism.” Nature (2010): 799-806.

15. Kenny, Tim. “Diabetes and High Blood Pressure.” Patient. Egton Medical Information Systems, 22 September. 2014. Web. 10 March. 2015. http://patient.info/health/diabetes-and-high-blood-pressure

16. “Understanding Blood Pressure Reasings.” American Heart Association. American Heart Association, 28 August. 2014. Web. 10 March. 2015. http://www.heart.org/HEARTORG/Conditions/HighBloodPressure/AboutHighBloodPressure/Understanding-Blood-Pressure-Readings_UCM_301764_Article.jsp

17. “The A1C Test and Diabetes.” National Diabetes Information Clearinghouse. National Institute of Digestive and Kidney Diseases, National Institutes of Health, 17 September. 2014. Web. 10 March. 2015. http://diabetes.niddk.nih.gov/dm/pubs/A1CTest/

18. “Guide to HbA1c.” Diabetes Co UK. 2015 Diabetes Digital Media Ltd. Web. 10 March. 2015. http://www.diabetes.co.uk/what-is-hba1c.html

19. Kern, P. A, G. B. Di Gregorio, T. Lu, N. Rassouli, and G. Ranganathan. “Adiponectin Expression From Human Adipose Tissue Relation to Obesity, Insulin Resistance, and Tumor Necrosis Factor-α Expression.” Diabetes (2003), 1779-1785.